SORT 41 (2) July-December 2017, 255-276 DOI: 10.2436/20.8080.02.59

A bivariate response model for studying the marks obtained in two jointly-dependent modules in higher education

Emilio Gómez-Déniz[1], Nancy Dávila Cárdenes[1] and María D. García Artiles[2]

Abstract

We study the factors which may affect students' marks in two modules, mathematics and statistics, taught consecutively in the first year of a Business Administration Studies degree course. For this purpose, we introduce a suitable bivariate regression model in which the dependent variables have bounded support and the marginal means are functions of explanatory variables. The marginal probability density functions have a classical beta distribution. Simulation experiments were performed to observe the behaviour of the maximum likelihood estimators. Comparisons with univariate beta regression models show the proposed bivariate regression model to be superior.

MSC: 62P25, 62E99, 97D60.

Keywords: Beta distribution, bivariate beta distribution, conditional distributions, covariate, marginal distributions, regression, mathematics, statistics, business studies.

1. Introduction

Event counts such as the number of claims for third-party liability, other claims under guarantee, medical consultations, the use of prescription drugs, and voluntary and/or involuntary job changes, among many others, are likely to be jointly dependent. In these cases, it is of interest to study how different covariates or factors may simultaneously affect the two random (dependent) variables involved. Bivariate Poisson regression models, bivariate negative binomial regression models (see Maher, 1990) and their extensions (see Gurmu and Elder, 2000), among other approaches, have been applied in these settings. Nevertheless, few such studies have been conducted when the dependent variables are continuous and bounded.

[1] Department of Quantitative Methods and T*i*DES Institute. University of Las Palmas de Gran Canaria, Spain. emilio.gomez-deniz@ulpgc.es, nancy.davila@ulpgc.es

[2] Department of Quantitative Methods. University of Las Palmas de Gran Canaria, Spain. mdgartiles@ulpgc.es
Received: March 2016
Accepted: March 2017

In the univariate case, Papke and Wooldridge (1996) examined potential econometric alternatives when the dependent variable is fractional, in a study of employee participation rates in 401(k) pension plans. More recently, Papke and Wooldridge (2008) analysed test pass rates and the portfolio choices of Australian households. Other research work related to the beta regression model includes Cepeda-Cuervo (2001), Paolino (2001), Ferrari and Cribari-Neto (2004) and Huang and Oosterlee (2011). The model proposed by Gómez-Déniz, Sordo and Calderín-Ojeda (2013) provides an alternative to the beta regression model, and affords a better fit, at least in an actuarial setting. The model proposed by Pérez-Rodríguez and Gómez-Déniz (2015) also appears to be comparable to the beta regression approach in financial econometrics. Using Bayesian methodology, Bayes, Bazán and García (2012) presented a variation of the beta regression model, while Cepeda-Cuervo and Núñez-Antón (2013), used spatial regression in an analysis of the quality of education. In the bivariate case, Cepeda-Cuervo, Achcarb and Garrido (2014) proposed a bivariate beta regression model with joint modelling of the mean and dispersion parameters.

As an extension of the works related above, we propose a flexible bivariate fractional response model in which the dependent variables are bounded and the marginal means are functions of explanatory variables.

Although a bivariate regression model could be built by using, for instance, copulas from the Sarmanov family of distributions (see Lee (1996)), we chose the bivariate beta regression proposed by Olkin and Liu (2003) for this study because it is a simple model with which to compute marginal distributions, means and variances. In this model, the beta distribution has a straightforward formulation in which the Euler Gamma function is the only one considered. In this respect, Cepeda-Cuervo et al. (2014) used copulas to obtain a bivariate beta regression model in which, as in our own model, the marginal distributions are beta. These authors assumed weak dependence between the variables of interest and modelled the dependence using a Farlie-Gumbel-Morgenstern copula function.

The model we propose is less complex than that presented in Cepeda-Cuervo et al. (2014) and therefore, by the Ockham's razor principle, it might be preferable (Jaynes, 1994).

In this paper, we study how some covariates may simultaneously affect the marks (ranging from 0 to 10) obtained by students in two first-degree subjects – Mathematics for Business and Basic Statistics in Business Administration Studies – taught at the University of Las Palmas de Gran Canaria (Spain) during two consecutive terms (first mathematics and then statistics). We assume that a good knowledge of mathematics will significantly influence the student's understanding of statistics and therefore that there exists a positive correlation between these two variables. Accordingly, the model proposed would be suitable for studying this relationship.

The importance of mathematical skills in other quantitative disciplines has been widely examined. In the fields of business and economics, many studies have analysed basic mathematical abilities as determinant factors of academic performance among first year university students: see, for example, Johnson and Kuennen (2006), Dolado and Morales (2009), Lunsford and Poplin (2011) and Arnold and Straten (2012).

Study plans in business and economics courses are organised in different ways, depending on the institution, but all have in common a requirement of basic mathematics to favour a better understanding of subjects that require this skill as a tool to develop more complex theories. In the present study, we focus on the above-mentioned mathematics and statistics modules, to determine whether certain common factors might explain the students' marks obtained in each subject.

The rest of this paper is structured as follows. Section 2 describes the bivariate model proposed by Olkin and Liu (2003), from which we derive the proposed bivariate regression model. This model and its parameters are studied in Section 3. The data are described in detail in Section 4. In Section 5 we fit the marginal beta regression models and the bivariate beta regression models, comparing the univariate and the bivariate models. Finally the results obtained and the main conclusions drawn are reported in Section 6.

2. Modelling bivariate marks

Although mathematics and statistics are known to be logical and effective means of solving certain problems, most Business Administration students, especially those in the first and second years of their degree courses, are not interested in these course subjects. Indeed, numerous students in this area of study present some form of rejection of mathematics and statistics. Nevertheless, our empirical evidence shows that the marks obtained by students in statistics are positively related to those achieved in mathematics. We assume this is because mathematics is an instrumental subject that influences the results achieved in statistics.

Let Y_1 and Y_2 be two random variables which represent the marks achieved in mathematics and statistics, respectively. To address the study goal presented in the introduction, and taking into account the above comments, we need a bivariate distribution that meets the following conditions:

a) The support of the distribution should be bounded, since the marks are usually restricted to a given interval.

b) The bivariate distribution should provide a dependence structure.

c) The correlation between the two random variables should be positive. That is, $\rho(Y_1,Y_2) > 0$.

d) Preferably, $\Pr(Y_2 > y_2 | Y_1 > y_1)$ should be a nondecreasing function in y_1 for all y_2. Thus, the higher the mathematics mark, the greater the probability of obtaining higher marks in statistics.

e) Because we wish to study the factors which may affect the marks obtained in the two courses, using a regression analysis, the marginal mean (the response variable) should be expressed as a function of the explanatory variables through a simple expression.

In this case, the standard beta distribution may be extended to the bivariate case. Many bivariate beta distributions have been derived from an application or as extensions to or generalisations of other well-known bivariate beta distributions. Since the latter are used in a wide variety of applications, the development and derivation of new bivariate beta distributions has been extensively studied. Nevertheless, few such distributions present these five features simultaneously. One, however, was proposed by Olkin and Liu (2003), with the following probability density function (pdf):

$$f(y_1,y_2)=\frac{y_1^{a_1-1}y_2^{a_2-1}(1-y_1)^{a_2+a_3-1}(1-y_2)^{a_1+a_3-1}}{B(a_1,a_2,a_3)(1-y_1y_2)^{a_1+a_2+a_3}}, \tag{1}$$

where $0<y_i<1$ $(i=1,2)$, $a_i>0$ $(i=1,2,3)$ and where $B(a_1,a_2,a_3)$ is given by $B(a_1,a_2,a_3)=\prod_{i=1}^{3}\Gamma(a_i)/\Gamma(\sum_{i=1}^{3}a_i)$, where $\Gamma(\cdot)$ is the Euler Gamma function. Henceforth, we use the expression $(Y_1,Y_2)\sim\mathscr{BB}(a_1,a_2,a_3)$ when the two random variables (Y_1,Y_2) fit the pdf (1).

The marginal distributions of Y_1 and Y_2 are beta distributions with parameters (a_1,a_3) and (a_2,a_3), respectively. Thus, the marginal means, the variances and the cross moment are given by

$$E(Y_i)=\frac{a_i}{a_i+a_3},\quad i=1,2, \tag{2}$$

$$var(Y_i)=\frac{a_ia_3}{(a_i+a_3)^2(a_i+a_3+1)},\quad i=1,2.$$

$$E(Y_1Y_2)=\frac{a_1a_2\Gamma(a_1+a_3)\Gamma(a_2+a_3)}{m\Gamma(a_3)\Gamma(m+1)}\,{}_3F_2(\{m_1,m_2,m\};\{m+1,m+1\};1), \tag{3}$$

where $m_i=a_i+1$ $(i=1,2)$, $m=a_1+a_2+a_3$ and ${}_3F_2$ is the generalised hypergeometric function. For details about this special function see, for instance, Gottschalk and Maslen (1988). This can be computed using the Mathematica package (see Wolfram (2003)). Using (2) and (3) we can obtain the covariance, $cov(Y_1,Y_2)$, and the correlation between Y_1 and Y_2, $\rho(Y_1,Y_2)$. For reasons of space, these large expressions are not shown here. Olkin and Liu (2003) showed that the correlation is always positive, with values in the interval $(0,1)$. The following result is obtained for the conditional distribution:

$$f(y_1|y_2)=\frac{y_1^{a_1-1}(1-y_1)^{a_2+a_3-1}(1-y_2)^{a_1}}{B(a_1,a_2+a_3)(1-y_1y_2)^{a_1+a_2+a_3}}, \tag{4}$$

$$f(y_2|y_1)=\frac{y_2^{a_2-1}(1-y_2)^{a_1+a_3-1}(1-y_1)^{a_2}}{B(a_2,a_1+a_3)(1-y_1y_2)^{a_1+a_2+a_3}}. \tag{5}$$

After some algebra, we derive the conditional mean obtained from (4) and (5). Thus

$$\begin{aligned} E(Y_1|Y_2=y_2) &= \frac{a_1}{a_1+a_2+a_3}\,{}_2F_1(1,a_2+a_3;a_1+a_2+a_3;y_2), \\ E(Y_2|Y_1=y_1) &= \frac{a_2}{a_1+a_2+a_3}\,{}_2F_1(1,a_1+a_3;a_1+a_2+a_3;y_1), \end{aligned} \tag{6}$$

where ${}_2F_1$ represents the hypergeometric function (see Gradshteyn and Ryzhik, 1994).

One of the advantages of using the pdf given in (1) is that for this distribution we have

$$\begin{aligned} \Pr(Y_2 > y_2|Y_1 > y_1^0) &\leq \Pr(Y_2 > y_2|Y_1 > y_1^1), \quad y_1^0 < y_1^1, \\ \Pr(Y_2 \leq y_2|Y_1 \leq y_1^0) &\geq \Pr(Y_2 \leq y_2|Y_1 \leq y_1^1), \quad y_1^0 < y_1^1, \end{aligned}$$

for all y_2. In other words, $\Pr(Y_2 > y_2|Y_1 > y_1)$ is a nondecreasing function in y_1 for all y_2 and $\Pr(Y_2 \leq y_2|Y_1 \leq y_1)$ is a nonincreasing function in y_1 for all y_2, because the pdf (1) is positively likelihood ratio dependent (see Tong (1980) and Olkin and Liu (2003) for details). This is corroborated by the fact that in our case the random variables Y_1 and Y_2 are positively quadrant dependent, a concept introduced by Lehmann (1996). Thus, we have

$$\begin{aligned} \Pr(Y_2 > y_2|Y_1 > y_1) &\geq \Pr(Y_2 > y_2)\Pr(Y_1 > y_1), \\ \Pr(Y_2 \leq y_2|Y_1 \leq y_1) &\geq \Pr(Y_2 \leq y_2)\Pr(Y_1 \leq y_1). \end{aligned}$$

A possible interpretation of the parameters of the distribution in (1) is this. Let W be a random variable measuring a student's lack of mathematics skills for use in subjects such as mathematics, statistics and physics. Empirical evidence shows that when Business Administration students are asked about their skills in mathematics and statistics, most of them acknowledge inadequacy in this field. Let U_i $(i=1,2)$ be the random variable representing the student's willingness to study these subjects i $(i=1,2)$. Assuming that W and U_i can take values in $(0,\infty)$, then the marks obtained in subject i can be represented by the random variables

$$Y_i = \frac{1}{1+W/U_i} = \frac{U_i}{U_i+W}, \quad i=1,2.$$

The gamma distribution provides a flexible representation of a variety of distribution shapes, by varying the shape parameter. Let us now assume that the random variables U_1, U_2 and W are independent and follow a standard gamma distribution with shape parameters a_1, a_2 and a_3, respectively. Then, the random variable (Y_1,Y_2) follows the distribution given in (1).

In conclusion, the pdf given in (1) seems to be a suitable distribution to model the joint random variables corresponding to mathematics and statistics marks when the latter are influenced by the former.

3. Regression model and estimation

Let us now consider a more realistic model, in which covariates are included. The linear regression model, which makes no distributional assumptions, is likely to be unsatisfactory because certain combinations of parameters and regressors could violate the nonnegative restriction and the upper limit on the mean. To avoid this situation we propose a parametric model based on using the distributional assumptions presented in the previous section.

When a regression analysis is to be performed, it is often useful to model the mean of the response. By equating the mean given in (2) to μ_i $(i = 1,2)$, solving for a_i $(i = 1,2)$, taking $a_3 = \theta$ and replacing the resulting expression in the pdf of the bivariate beta distribution in (1), we obtain the following reparametrisation.

$$f(y_1,y_2) = \frac{y_1^{\phi_1\mu_1-1} y_2^{\phi_2\mu_2-1}(1-y_1)^{\phi_2-1}(1-y_2)^{\phi_1-1}}{B(\phi_1\mu_1,\phi_2\mu_2,\theta)\,(1-y_1y_2)^{(1-\mu_1\mu_2)\phi_1\phi_2/\theta}}, \tag{7}$$

where $\phi_i = \theta/(1-\mu_i)$, $0 < \mu_i < 1$, $i = 1,2$; with $0 < y_1 < 1$, $0 < y_2 < 1$ and $\theta > 0$. Under this reparametrisation of the bivariate beta distribution, the marginal mean is $E(Y_i) = \mu_i$, for $i = 1,2$.

Now, let $\boldsymbol{x}_{\kappa i}^{\mathsf{T}} = (x_{1i}, x_{2i}, \ldots, x_{pi})$ be a vector of the p covariates associated with the ith observation. This is a vector of linearly independent regressors that are thought to determine (y_1, y_2). For the ith observation, the model takes the form

$$(Y_{1i}, Y_{2i}) \sim \mathscr{B}\mathscr{B}(\mu_{1i}, \mu_{2i}, \theta),$$

$$\mu_{\kappa i}(\boldsymbol{x}_{\kappa i}, \boldsymbol{\beta}_\kappa) \equiv \mu_{\kappa i} = \frac{\exp(\boldsymbol{x}_{\kappa i}^{\mathsf{T}}\boldsymbol{\beta}_\kappa)}{1+\exp(\boldsymbol{x}_{\kappa i}^{\mathsf{T}}\boldsymbol{\beta}_\kappa)}, \quad \kappa = 1,2.$$

Here, $i = 1,\ldots,n$ denotes the number of observations, $\boldsymbol{x}_{\kappa i}$ denotes a vector of p explanatory variables for the ith observation and $\boldsymbol{\beta}_\kappa = (\beta_{\kappa 1},\ldots,\beta_{\kappa p})^{\mathsf{T}}$, $\kappa = 1,2$, denotes the corresponding vectors of regression coefficients. It is clear that each variable Y_1 and Y_2 may be influenced by different characteristics and variables. For this reason, the explanatory variables that are used to model each mean $\mu_{\kappa i}$, may not be the same. Furthermore, observe that the logit link assumed ensures that $\mu_{\kappa i}$ falls within the interval $(0,1)$.

Under this model the log-likelihood function takes the form given in the Appendix, which shows the equations used to provide the estimates of the parameters. The above model presents the advantage of simplicity; on the other hand, the normal equations require the use of the digamma function, $\psi(z) = \frac{d}{dz}\log(\Gamma(z))$, $z > 0$, in order to estimate all the model parameters. However, this problem is overcome by means of Mathematica routines (see Wolfram, 2003) and RATS (see Brooks, 2009), which work well with this special function.

Because the equations which provide the estimates of the parameters cannot be solved explicitly, they must be addressed either by numerical methods or by directly maximising the log-likelihood function; in this study, the latter approach is adopted. Since the global maximum of the log-likelihood surface is not guaranteed, different initial values of the parametric space can be considered as seed points. In this sense, we have used the `FindMaximum` function of the Mathematica software package v.11.0 (Wolfram, 2003). Moreover, other methods provided by Mathematica, such as Newton, PrincipalAxis and QuasiNewton (all of which are available in Mathematica) obtain the same result. Finally, the standard errors of the estimated parameter are approximated by inverting the Hessian matrix. This can also be done by approximating the Hessian matrix and recovering it from the Cholesky factors. These parameters were also computed by the RATS package, and the same values were obtained.

3.1. Marginal effects

The marginal effect reflects the variation of the conditional mean produced by a one-unit change in the jth covariate $(j = 1,\ldots,p)$. The marginal effect can be calculated as $\delta_j = \frac{\partial \mu_{\kappa i}}{\partial x_{ji}} = \beta_{\kappa j}\mu_{\kappa i}(1-\mu_{\kappa i})$, $\kappa = 1,2$; $i = 1,\ldots,n$; $j = 1,\ldots,p$. Thus, the marginal effect indicates that a one-unit change in the jth regressor increases or decreases the expectation of marks for the jth covariate by δ_j units, $j = 1,\ldots,p$. This expression is the one normally obtained under the logit marginal effect. For indicator variables which takes only the values 0 or 1 the marginal effect is $\delta_j = E(y_\kappa|x_{ji} = 1)/E(y_\kappa|x_{ji} = 0) \approx \exp(\beta_{\kappa j})$, $\kappa = 1,2$; $i = 1,\ldots,n$; $j = 1,\ldots,p$. Therefore, the conditional mean is $\exp(\beta_{\kappa j})$ times larger if the indicator variable is one rather than zero.

3.2. Simulation study

We now present some simulation results, obtained by a bootstrap experiment, to study the behaviour of the maximum likelihood estimators. The Mathematica package was used to create random variables from the pdf (7). In this process, the first component of the vector was generated from a marginal, and then a second one from a conditional distribution. The estimated values of the parameters were then computed directly using the `FindMaximum` function of Mathematica v.11.0 (Wolfram (2003)). The following sets of model parameters were considered:

Table 1: *Average estimates (first row), the square root of the mean squared errors (second row in parenthesis) and the correlation (ρ) between estimated parameters based on 1000 replications.*

n	$\mu_1 = 0.15$	$\mu_2 = 0.25$	$\theta = 0.85$	$\rho(\mu_1,\mu_2)$	$\rho(\mu_1,\theta)$	$\rho(\mu_2,\theta)$
25	0.1402	0.2213	0.9827	0.7524	–0.7481	–0.6804
	(0.0367)	(0.0585)	(0.4490)			
50	0.1932	0.3004	0.7256	0.6425	–0.8068	–0.7392
	(0.0303)	(0.0438)	(0.1364)			
75	0.1795	0.2537	0.7326	0.6009	–0.7757	–0.7566
	(0.0246)	(0.0310)	(0.1116)			
100	0.1526	0.2820	0.7292	0.4851	–0.7637	–0.5173
	(0.0167)	(0.0269)	(0.0842)			
	$\mu_1 = 0.25$	$\mu_2 = 0.75$	$\theta = 0.5$	$\rho(\mu_1,\mu_2)$	$\rho(\mu_1,\theta)$	$\rho(\mu_2,\theta)$
25	0.2308	0.7273	0.6525	0.4976	–0.7454	–0.5647
	(0.0558)	(0.0347)	(0.1829)			
50	0.2542	0.7529	0.4906	0.5719	–0.7883	–0.5647
	(0.0375)	(0.0323)	(0.0998)			
75	0.2395	0.7190	0.5886	0.4495	–0.7168	–0.3814
	(0.0326)	(0.0274)	(0.0715)			
100	0.2992	0.7943	0.4170	0.4116	–0.7040	–0.3814
	(0.0298)	(0.0180)	(0.0438)			
	$\mu_1 = 0.50$	$\mu_2 = 0.15$	$\theta = 1.50$	$\rho(\mu_1,\mu_2)$	$\rho(\mu_1,\theta)$	$\rho(\mu_2,\theta)$
25	0.5210	0.1273	1.9765	0.1815	0.0217	–0.7145
	(0.0484)	(0.0268)	(0.4869)			
50	0.5713	0.1603	1.4338	0.4822	–0.4643	–0.5710
	(0.0332)	(0.0249)	(0.1888)			
75	0.5289	0.1600	1.3432	0.1993	–0.4655	–0.5616
	(0.0293)	(0.0199)	(0.1618)			
100	0.5063	0.1851	1.5894	0.5823	–0.6167	–0.6597
	(0.0240)	(0.0198)	(0.2168)			

$$(\mu_1,\mu_2,\theta) = (0.15, 0.25, 0.85),$$
$$(\mu_1,\mu_2,\theta) = (0.25, 0.75, 0.50),$$
$$(\mu_1,\mu_2,\theta) = (0.50, 0.15, 1.50).$$

In all three cases, we have simulated observations with a sample size given by $n = 25, 50, 75$ and 100. We report the average estimates and the square root of the mean squared errors based on 1000 replications, i.e. the bootstrap sample is taken from the original by using sampling with replacement 1000 times. Additional replications are considered unnecessary, as the computational time needed would be prohibitive; nevertheless, we acknowledge that the use of fewer replications might reduce the statistical accuracy obtained. The results are shown in Table 1. In general, as the sample size increases the estimates approach the true values and the biases and the mean squared er-

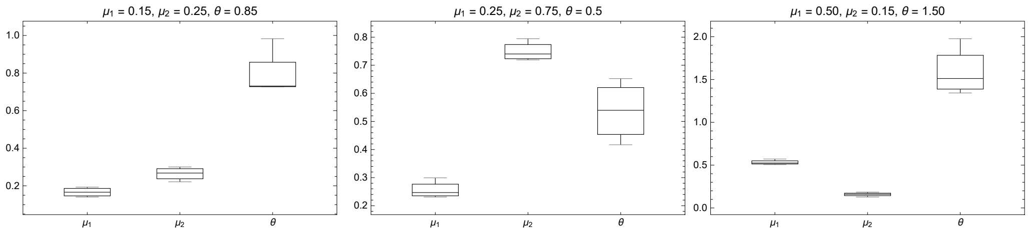

Figure 1: *Box-and-whisker charts showing the differences between the true parameter values and the estimates based on the data in Table 1.*

rors decrease. These outcomes corroborate the consistency of the maximum likelihood estimates. From the standard errors obtained, it is evident that the errors are smaller as the sample size increases. Furthermore, the correlation between the parameters is always positive for μ_1 and μ_2 and negative for μ_2 and θ. Hence, the correlation between these two sets of parameters is not very high. Finally, Figure 1 shows that the parameters estimated have a slight negative bias, which is more apparent in the θ parameter.

4. Factors affecting the mathematics and statistics marks obtained

In order to make use of the bivariate regression model, we examined the relation between the marks achieved by the students in two course subjects: Mathematics and Statistics in Business Administration. Most of these students, before entering the university had studied subjects focused on statistics, more so than basic mathematics. In fact, many of them believed they did not need mathematics and did not consider the two courses to be related. During the first term, difficulties were encountered in mathematics, but with the start of the statistics class, in the second term, the students believed their performance would be better. Therefore, at the beginning of the mathematics course, the students were informed of the analysis that would be conducted, and were asked to complete a questionnaire on this subject. The following section describes how the data were compiled and how many students comprised the final study sample.

4.1. The sample

The data for this study were collected from eight student groups in the Mathematics for Business and Statistics modules taught during the first year of the Business Administration degree course at the University of Las Palmas de Gran Canaria (Spain). The study population was initially composed of 725 students enrolled in these groups. On the first day of classes in 2013, a questionnaire was handed out to 456 students.

The final sample was composed of the 213 students who completed both modules (mathematics and statistics) and answered the survey. The questionnaire was divided

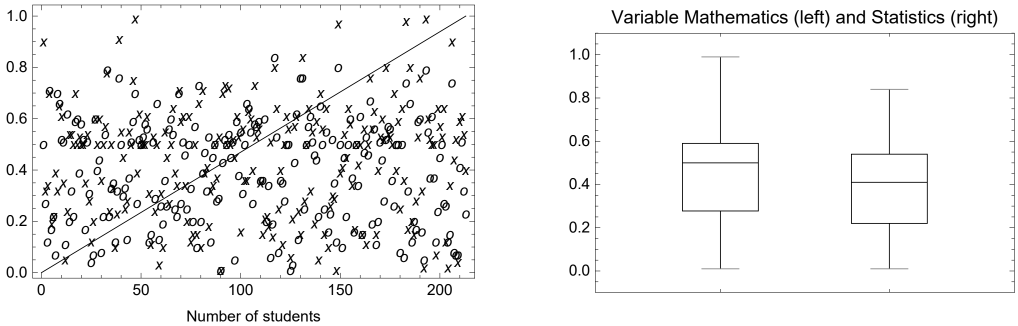

Figure 2: *Scatter plot of the marks in mathematics (×) and statistics (○) on the left and box-and-whisker charts for the two dependent variables on the right.*

into two parts; the first contained questions dealing with personal and academic information, and the second presented four short mathematics exercises.

During the academic year, the students are given three opportunities to take the class exams. In total, 114 students (54%) of the students who completed the initial questionnaire passed the Mathematics for Business exam, and 92 (43%) passed in Basic Statistics. The final marks for the students in the sample, for each of these two subjects, are shown in the scatter plot in Figure 2. Few students obtained high marks in mathematics, and there was a large concentration of values below 0.6. Figure 2 also shows box-and-whisker charts for the two dependent variables. Since the support of (1) does not include the values zero or one, and taking into account that the data contained very few such marks, instead of removing them, these marks were replaced by 0.001 and 0.999, respectively.

The range of possible exam marks was from 0 to 10. In Spain, a pass mark is 5 or more. In the study sample, most of the marks obtained were between 4.5 and 7.0. Many were under 4.0 and very few were over 7.0.

4.2. Personal and academic factors

The survey data collected concerning personal and academic factors are shown in Table 2. Among the personal and academic information sought in the questionnaire, the variable AGE was obtained by dividing the students into those born in 1995 and those born earlier. The year 1995 was taken because in 2013, when the study data were compiled, these students would be aged 18 years, which is the usual age for university entrance.

To accompany the continuous variable AGE, a box-and-whisker chart is included in Figure 3. This chart shows that most of the students in the study sample were aged 18-21 years. The outliers in the sample were aged 27-32 years. These students had entered the university at age 25 years or older, and had had to pass specific examinations to do so.

Table 2: *Descriptive data: personal and academic information.*

Personal and academic information		% Respondents N = 213	% Mathematics Passes N = 114	% Statistics Passes N = 92
AGE	Born in 1995	47	43	40
	Born before 1995	53	57	60
SCHOOL	Public	82	80	84
	Private	18	20	16
TRACK	Technical-Science	12	14	12
	Other	88	86	88
ADMSCORE	$[5, 8]$	54	45	49
	$(8, 14]$	46	55	51
PREF	Yes	90	94	92
Business	No	10	6	8
	Yes	71	68	64
NEWCOMER	No	29	32	36
GRANT	Yes	65	63	62
	No	35	37	38
WORKING	Yes	4	4	8
	No	96	96	92
GENDER	Male	48	51	51
	Female	52	49	49

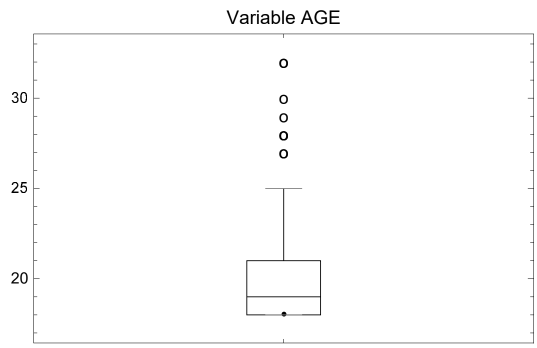

Figure 3: *Box-and-whisker chart of the continuous variable AGE.*

Other information requested concerned the type of school (public or private) attended before university entrance. This variable was termed SCHOOL, and the academic specialisation chosen by the student during the last two years of high school was termed TRACK. Different types of track are available, but for the purposes of this study, they were divided into Technical-Science and Others.

In Spain, university entrance requires a specific examination, known as PAU (Prueba de Acceso a la Universidad) to be taken, in addition to the final high-school exams.

The weighted average of the latter mark and the PAU result determines the final admission score obtained; we term this variable ADMSCORE. The threshold for university entrance ranges from 5 to 14, depending on demand for the course and on the places available. The students in our sample were divided into those who obtained an admission score of 8 or less, and those who obtained 9 or more. After taking the PAU exam, a period is allowed during which students may choose the degree course they wish to study. For the purposes of this research, to identify the strength of vocation in the students' choice, the questionnaire asked whether the degree in Business was their first preference. The study variable in this respect was termed PREF.

Every year, 400 new students enrol in the Mathematics for Business module. However, the total number of students enrolled each year is almost double this figure due to the high number of students who failed to pass or did not sit the previous year's exam, and who had to retake the course. To distinguish new students from those retaking the course, this information was requested, and the corresponding variable was termed NEWCOMER. Other variables included were the students' gender (GENDER), whether they were receiving a study grant (GRANT) and whether they were working (WORKING).

4.3. Factors related to mathematics skills

The variables concerning the students' mathematics skills are shown in Table 3. As part of the questionnaire, the students were asked to solve four exercises and to describe how they had done so. Different steps were involved in each exercise. The score awarded for

Table 3: *Descriptive data for mathematics skills.*

BASIC MATHEMATICS SKILLS	Variables	Exercices	% Respondents	% Pass rate
1. LINEAR EQUATION				
Handle rational coefficients	FRACTIONS	$\frac{1}{2}x+\frac{3}{4}x=0$	46	55
Solving the equation	LINEAREQ		33	38
2. EQUATIONS SYSTEM				
Resolution system method	SYSTEM		64	73
Raising the quadratic equation	EQ2	$\left.\begin{aligned} -x^2+2x-3y&=0,\\ 5x+3y&=0.\end{aligned}\right\}$	48	52
Solving the quadratic equation	SOLVINGEQ2		22	21
Discuss the solutions	DISCUSSEQ2		10	12
3. ALGEBRAIC EXPRESSIONS				
Clear the unknown	CLEANUNKN	$2x^2y^3-ax^3y=0$	24	31
Simplify exponents	SIMPLIFYEXP		17	25
4. BASIC DERIVATIVES				
Polynomial with integer exponent	DERINTEXP		47	53
Polynomial with rational exponent	DERRATEXP	$f(x)=\frac{1}{3}x^3-x^{1/2}+2$	23	34
Simplify the final expression	SIMPLYFYDER		11	18

each step in the exercise was 1 when it was performed correctly, or 0 otherwise. The following variables were associated with each step in the procedure.

The first exercise was a linear equation with rational coefficients. The question was evaluated according to whether the student was able to handle the basic algebra involved in manipulating the coefficients. The study variable in this respect was termed FRACTIONS. Solving the equation and obtaining the value of the unknown parameter were represented by the term LINEAREQ.

The second question concerned a very basic non-linear equation system. The marking criteria were defined using the following variables: the application of a method for solving linear systems, SYSTEM; giving the resulting incomplete quadratic equation, EQ2; solving the equation and discussing the solution obtained, SOLVINGEQ2 and DISCUSSEQ2, respectively.

The third exercise consisted in giving the value of the parameter "*a*" after simplifying the algebraic expression (see Table 3). To achieve a positive score, the students had to clear the unknown, CLEANUNKN, and simplify the exponents, SIMPLIFYEXP. In the final exercise, the students were asked to calculate the derivative function of an elementary polynomial expression including an integer and rational exponents. The exercise was evaluated according to whether the integer exponent was correctly derived, DERINTEXP, whether the derivative of the rational expression was correctly given, DERRATEXP, and whether the last expression was correctly simplified, SIMPLYFYDER.

Observation of the final column in Table 3, the percentage of students who passed the final subject exams, clearly shows that although the percentage of passing students is higher among those who correctly responded to the questions in the initial survey, the pass rates are still unacceptably low. The results obtained reflect a lack of basic skills in some areas of mathematics.

5. Testing the models

The descriptive values obtained for the dependent variables are given by $E(Y_1) = 0.44169$, $var(Y_1) = 0.04728$, $E(Y_2) = 0.39145$ and $var(Y_2) = 0.04000$. Thus, the mean value of Y_1 is larger than that of Y_2 while the variance is similar in each case. The correlation is positive, with a value of 0.67243 indicating that these values increase or decrease together.

5.1. Model without covariates

The model was initially implemented without covariates, which produced the parameter values shown in Table 4. The standard errors are shown between parentheses, and the estimates obtained when the univariate beta distribution is assumed are also shown. In addition, we show the value obtained for the Akaike information criterion (AIC).

Table 4: *Univariate and bivariate models without covariates.*

	Y_1 : Maths	Y_2 : Statistics	(Y_1, Y_2)
$\widehat{\mu}_1$	0.441821		0.438451
	(0.015631)		(0.014354)
$\widehat{\mu}_2$		0.384893	0.388744
		(0.013748)	(0.014500)
$\widehat{\theta}$	3.561590	4.745440	2.497550
	(0.307957)	(0.422396)	(0.179619)
AIC	–40.832	–93.8506	–203.260

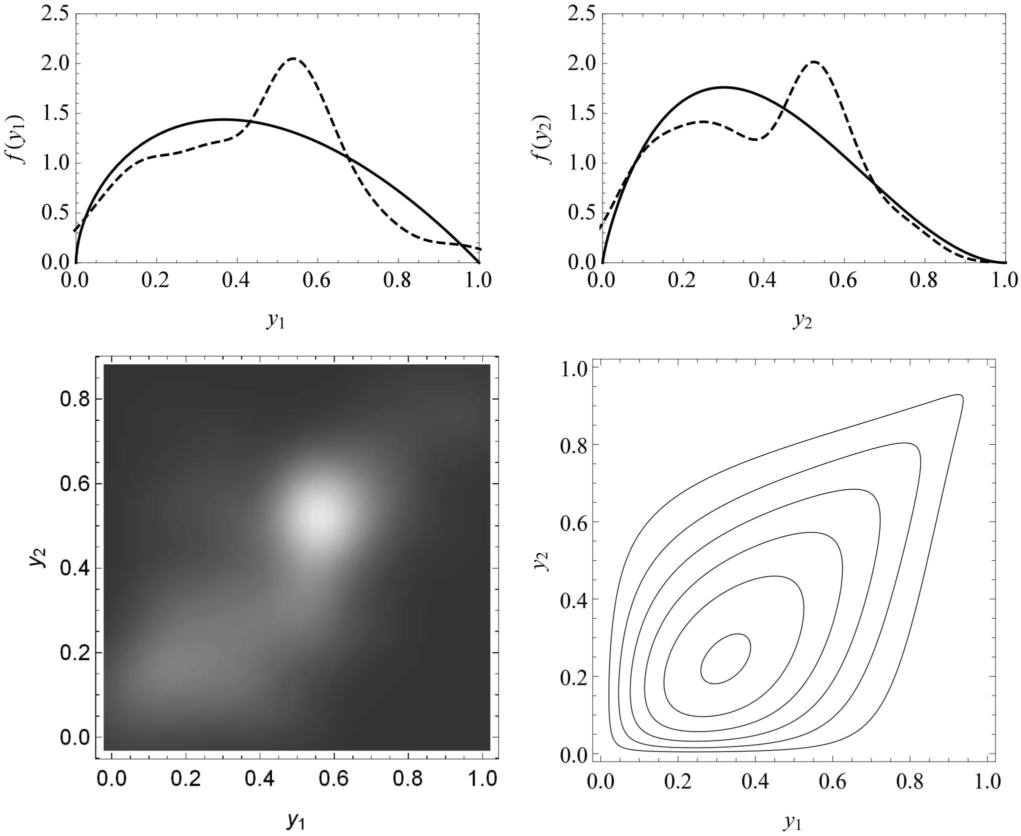

Figure 4: *Top: the smooth kernel densities (dashed curves) and the pdfs (solid curves) of the estimated univariate beta distribution. Bottom: the smooth contour plot obtained from the data (left) and the estimated contour plots of the bivariate beta pdf.*

(AIC $= 2(k - \ell_{\max})$, where k is the number of model parameters and $\ell_{\max}$ is the maximum value of the log-likelihood function; see Akaike (1974) for details.) These values should be compared only with the fitted models obtained when covariates are included, as described below.

From the parameter values obtained by the model without covariates, we obtain the estimated descriptive statistics $\widehat{E(Y_1)} = 0.4384$, $\widehat{var(Y_1)} = 0.0452$, $\widehat{E(Y_2)} = 0.389$ and $\widehat{var(Y_2)} = 0.047$. The estimated correlation result is 0.417. Thus, except for the correlation, the estimated values are close to the true empirical values.

Figure 4 shows (top) the smooth kernel density and the pdf of the estimated univariate beta distribution. The lower part of the figure shows the smooth contour plot obtained from the data (left) and the estimated contour plots of the bivariate beta pdf. Clearly, the univariate beta distribution provides a better fit to the sample values of the mathematics marks than to those for statistics. The contour plot has a similar shape to the smooth contour plot, from which we conclude that the bivariate distribution is a better model than the univariate one.

5.2. Including covariates

The two models, univariate and bivariate, were then evaluated, making use of all the covariates described in Tables 2 and 3. The normal equations and Fisher's information matrix for the univariate beta regression model, given by

$$f(y_i) = \frac{\Gamma(\theta)}{\Gamma(\theta\mu_i)\Gamma((1-\mu_i)\theta)} y_i^{\mu_i\phi_i - 1}(1-y_i)^{(1-\mu_i)\theta-1}, \quad i = 1,2$$

are discussed in detail in Ferrari and Cribari-Neto (2004). The regression results are shown in Tables 5 and 6. Better results are obtained with the bivariate regression model than by the separate estimation of two univariate beta regressions. When the covariance of the joint model is close to zero, the two models are nested. A likelihood-ratio test, comparing the bivariate value to the sum of the log-likelihood values of the separate estimation, provided further evidence of the advantages of the bivariate beta regression model.

The univariate model was then analysed for each module (see Table 5). With respect to personal and academic information, the following significant variables were obtained: AGE and ADMSCORE. In addition, the AGE-SQUARED variable was introduced to determine whether increased age was associated with poorer performance in this academic area.

In the univariate model, the marks for statistics did not seem to be influenced by the students' skills in mathematics, as the significant variables for statistics did not differ from those found for the Mathematics for Business class. This fact might be related to the students' background and/or to the class content. A good command of systems of equations, together with an understanding of derivatives and of the simplification process can have a positive effect on the marks obtained for mathematics, because a large amount of basic calculus is included in the topics addressed in this subject.

The bivariate model obtains better results because new significant variables are present. On the one hand, for personal and academic information, the significant variables are the same as in the univariate model, AGE and ADMSCORE (see Table 6). In both cases, the marginal effect is positive. Thus, the older the students and the better their admission score, the higher the marks obtained in the Mathematics for Business course. However, for the AGE-SQUARED variable, the marginal effect is negative. The age factor may have a positive effect on the students who are retaking the course, due to the knowledge acquired from the previous year, in the case of those whose age is close to that of the non-retakers (i.e. 18 years). On the other hand, when the AGE-SQUARED variable is considered, the students' additional age has a negative effect. We believe this is because older students have much greater difficulty in understanding the course contents. The same effects of the covariates were observed with respect to the statistics course. In the latter case, however, a further variable, NEWCOMER, was significantly present in the bivariate model, with a negative marginal effect. It may be relevant that the new students, before starting university studies, took a course focused on statistics, although not on calculus; however, this background does not seem to have any positive impact on their later performance.

Table 5: *Details for univariate fitted models including covariates.*

		MATHEMATICS		
Personal and academic information				
Variable	Coeff	Std Error	$\|t\|$-Stat	p-value
AGE	0.47428300	0.24773300	1.91450000	0.05695610
AGE-SQUARED	–0.00859712	0.00535907	1.60422000	0.11021300
ADMSCORE	0.47627700	0.12479800	3.81639000	0.00017952
Skill in Mathematics				
Variable	Coeff	Std Error	$\|t\|$-Stat	p-value
SYSTEM	0.35490300	0.12840300	2.76398000	0.00623282
DERRATEXP	0.46429600	0.18433700	2.51874000	0.01254520
SIMPLIFYDER	0.58471300	0.24651500	2.37192000	0.01862630
		STATISTICS		
Personal and academic information				
Variable	Coeff	Std Error	$\|t\|$-Stat	p-value
AGE	0.31052900	0.23665300	1.31217000	0.19091700
AGE–SQUARED	–0.00528830	0.00513981	1.02889000	0.30473200
ADMSCORE	0.34814600	0.12009600	2.89889000	0.00414850
Other parameters				
Constant for Mathematics	–6.80259000	2.81041000	2.42050000	0.01637510
Constant for Statistics	–4.65696000	2.66416000	1.74800000	0.08194620
θ for Mathematics	4.88422000	0.43697500	11.17730000	0.00000000
θ for Statistics	5.06500000	0.45354600	11.16760000	0.00000000
Value of the AIC for Mathematics: –94.917				
Value of the AIC for Statistics: –101.721				

Table 6: *Details for bivariate fitted model including covariates.*

		MATHEMATICS			
Personal and academic information					
Variable	Coeff	Std Error	$\|t\|$-Stat	p-value	δ_j
AGE	0.228201842	0.183687046	1.24234	0.21411099	1.256
AGE-SQUARED	–0.003363921	0.004001199	0.84073	0.40050008	0.996
ADMSCORE	0.434494686	0.090404187	4.80613	0.00000154	1.544
Skill in Mathematics					
Variable	Coeff	Std Error	$\|t\|$-Stat	p-value	δ_j
FRACTIONS	0.254534936	0.089643529	2.83941	0.00451967	1.289
SYSTEM	0.412448241	0.124542939	3.31170	0.00092733	1.510
EQ2	–0.365876393	0.124355071	2.94219	0.00325899	0.693
SIMPLIFYEXP	0.421013978	0.110824320	3.79893	0.00014532	1.523
DERRATEXP	0.299258577	0.130434786	2.29432	0.02177237	1.348
SIMPLIFYDER	0.434442994	0.159227590	2.72844	0.00636346	1.544
		STATISTICS			
Personal and academic information					
Variable	Coeff	Std Error	$\|t\|$-Stat	p-value	δ_j
AGE	0.152238309	0.210965922	0.72163	0.47052499	1.164
AGE-SQUARED	–0.001916838	0.004500367	–0.42593	0.67015944	0.998
ADMSCORE	0.375708496	0.099697582	3.76848	0.00016424	1.456
FRESHMEN	–0.357923647	0.108415456	3.30141	0.00096201	0.700
Skill in Mathematics					
Variable	Coeff	Std Error	$\|t\|$-Stat	p-value	δ_j
DERINTEXP	0.205826943	0.089275677	2.30552	0.02113741	1.228
Other parameters					
θ	3.164211826	0.199656264	15.84830	0.00000000	
Constant for Mathematics	–3.989088242	2.067120538	1.929780	0.05363408	
Constant for Statistics	–2.720631590	2.433508707	1.117990	0.26357246	
Value of the AIC: –285.820					

In the bivariate model, with respect to mathematics skills, some of the variables observed in the univariate model were again found to be relevant; in addition, the folowing new ones appeared: FRACTIONS, SYSTEM, EQ2, SIMPLIFYEXP, DERRATEXP and SIMPLIFYDER. In every case, the marginal effects were positive. Thus, when students are competent with the basic algebra of rational expressions, they are more likely to obtain higher marks in mathematics. The same is true when they can correctly apply a method for resolving a linear equations system to generate a quadratic equation to be solved. Another factor that appears to be significant is the ability to simplify algebraic expressions, to derive polynomial functions with rational exponents and to simplify the expression of the derivative function obtained. For these covariates, the positive marginal effects mean that the students' marks increase when they are able to correctly complete the exercises in question. However, the corresponding results for the statistics course show that the only significant factor was the covariate defining whether the students were capable of determining the derivative of a polynomial expression with integer exponents. Success in this task was also associated with higher marks in the subject, possibly because this type of expression appears in some elements of the statistics course.

6. Results and conclusions

As part of the Business Administration degree offered by the University of Las Palmas de Gran Canaria (Spain), a Mathematics for Business course is taught in the first term of the first year; this is followed by a course focusing on applied statistics in social sciences. In view of the obvious connection between these two courses, we decided to analyse the relationship between the marks obtained in each course and to determine which covariates might affect these marks.

Accordingly, we considered a flexible bivariate regression model to be applied when the dependent variables are bounded and the marginal means are functions of the explanatory variables. This model was applied to study the personal and academic factors relevant to the students in our study sample and the basic mathematical skills that may affect the marks obtained in the above-mentioned courses (mathematics and statistics). In our opinion, the model proposed is competitive with that presented by Cepeda-Cuervo et al. (2014), who generated a bivariate beta regression model from copulas evaluating it using a Bayesian methodology. As in our own case, the marginal distributions of the latter model were beta, but these authors assumed a weak dependence between the variables of interest, which was modelled by a Farlie-Gumbel-Morgenstern copula function. The model we propose has fewer parameters and therefore is simpler.

The results obtained in the present analysis show that the mean value of the marks obtained increases with the age of the students, in both courses. Specifically, the students who were born before 1995 had higher marks both for mathematics and for statistics. We interpret this finding as follows: some of the students in the final sample had been enrolled in the same mathematics course the previous year, and so they were not newcomers to the subject. Indeed, some had taken remedial courses, or had transferred from other undergraduate studies. Thus, following an initial lack of success, these students subsequently acquired mathematics skills enabling them to achieve better marks in the subject.

With regard to the admission score variable, this too was significant for both subjects, with a positive marginal effect. Thus, the higher the admission mark the better the marks obtained for mathematics and statistics. In this respect, obviously, the best students were most likely to achieve the best marks in mathematics and statistics.

Among the other variables related to personal information, another relevant factor was whether the students were newcomers, i.e. studying these subjects for the first time. Nevertheless, this variable was only significant for the statistics subject, which probably reflects the background acquired in this respect in the Social Sciences track studied at high school.

Finally, with regard to the influence of mathematical skills on the marks obtained for statistics, only the variable related to obtaining the derivative of polynomial expressions with integer exponents was found to be significant. It is striking that no other mathematical ability affected the marks for statistics. This might be because the basic statistics course in question is mainly descriptive, merely introducing the main concepts; consequently, most of the students were already acquainted with these concepts having opted for the Social Science track at high school. Despite these considerations, however, the marks obtained for statistics and the success rate in this course were even worse than for the business mathematics course.

In the light of the results obtained, we conclude that the bivariate beta regression model is more suitable than the univariate model for the analysis described in this paper.

Acknowledgements

The authors would like to express their gratitude to the Associate Editor and to an anonymous reviewer for their valuable comments and suggestions. EGD was partially funded by grant ECO2013-47092 (Ministry of Economy and Competitiveness, Spain).

Appendix

We present the equations needed to perform the estimation using the maximum likelihood method when covariates are introduced into the model. Consider a sample consisting of n observations $(\tilde{y}_1,\tilde{y}_2) = \{(y_{11},y_{21}),\ldots,(y_{1n},y_{2n})\}$, taken from the probability function (7). The log-likelihood is given by

$$\begin{aligned}\ell \equiv \ell(\theta,\beta_1,\beta_2;(\tilde{y}_1,\tilde{y}_2)) = \sum_{i=1}^{n} &[(\phi_{2i}-1)\log(1-y_{1i}) + (\phi_{1i}-1)\log(1-y_{2i}) \\ &+(\phi_{1i}\mu_{1i}-1)\log y_{1i} + (\phi_{2i}\mu_{2i}-1)\log y_{2i} \\ &-\frac{\phi_{1i}\phi_{2i}}{\theta}(1-\mu_{1i}\mu_{2i})\log(1-y_{1i}y_{2i}) \\ &-\log B(\phi_{1i}\mu_{1i},\phi_{2i}\mu_{2i},\theta)], \end{aligned} \tag{8}$$

where $\phi_{\kappa i} = \theta/(1-\mu_{\kappa i})$, $\kappa = 1,2$.

From straightforward computation, we have

$$\frac{\partial \mu_{\kappa i}}{\partial \beta_{\kappa j}} = \mu_{\kappa i} x_{\kappa j}, \;\; \frac{\partial \phi_{\kappa i}}{\partial \beta_{\kappa j}} = \frac{1}{\theta}(\phi_{\kappa i}\mu_{\kappa i} x_{\kappa j})^2,$$

from which we obtain the first partial derivatives of the log-likelihood function (8) with respect to θ and $\beta_{\kappa j}$ ($\kappa = 1,2,\; j = 1,\ldots,p$), given by

$$\begin{aligned}
\frac{\partial \ell}{\partial \theta} &= \sum_{i=1}^{n} \left[\frac{\log(1-y_{1i})}{1-\mu_{2i}} + \frac{\log(1-y_{2i})}{1-\mu_{2i}} + \frac{\mu_{1i}\log y_{1i}}{1-\mu_{1i}} + \frac{\mu_{2i}\log y_{2i}}{1-\mu_{2i}} \right. \\
&\quad \left. + \frac{\phi_{1i}\phi_{2i}}{\theta^2}(1-\mu_{1i}\mu_{2i})\log(1-y_{1i}y_{2i}) - \frac{\psi(\theta)-\psi(\theta+\sum_{\kappa=1}^{2}\phi_{\kappa i}\mu_{\kappa i})}{B(\phi_{1i}\mu_{1i},\phi_{2i}\mu_{2i},\theta)} \right], \\
\frac{\partial \ell}{\partial \beta_{1j}} &= \sum_{i=1}^{n} \mu_{1i}x_{1j} \left[\left(\mu_{2i} - \frac{\phi_{1i}\mu_{1i}x_{1j}}{\theta}(1-\mu_{1i}\mu_{2i}) \right) \frac{\phi_{1i}\phi_{2i}}{\theta} \log(1-y_{1i}y_{2i}) \right. \\
&\quad + \left(1 + \frac{\phi_{1i}}{\theta}\mu_{1i}^2 x_{1j} \right) \phi_{1i}\log y_{1i} + \frac{\phi_{1i}^2\mu_{1i}x_{1j}}{\theta}\log(1-y_{2i}) \\
&\quad \left. + \frac{\mu_{1i}\phi_{1i}x_{1j}}{B(\phi_{1i}\mu_{1i},\phi_{2i}\mu_{2i},\theta)} \left(\psi(\mu_{1i}\phi_{1i}) - \psi(\theta + \sum_{\kappa=1}^{2}\phi_{\kappa i}\mu_{\kappa i}) \right) \right], \\
\frac{\partial \ell}{\partial \beta_{2j}} &= \sum_{i=1}^{n} \mu_{2i}x_{2j} \left[\left(\mu_{1i} - \frac{\phi_{2i}\mu_{2i}x_{2j}}{\theta}(1-\mu_{1i}\mu_{2i}) \right) \frac{\phi_{1i}\phi_{2i}}{\theta} \log(1-y_{1i}y_{2i}) \right. \\
&\quad + \left(1 + \frac{\phi_{2i}}{\theta}\mu_{2i}^2 x_{2j} \right) \phi_{2i}\log y_{2i} + \frac{\phi_{2i}^2\mu_{2i}x_{2j}}{\theta}\log(1-y_{1i}) \\
&\quad \left. + \frac{\mu_{2i}\phi_{2i}x_{2j}}{B(\phi_{1i}\mu_{1i},\phi_{2i}\mu_{2i},\theta)} \left(\psi(\mu_{2i}\phi_{2i}) - \psi(\theta + \sum_{\kappa=1}^{2}\phi_{\kappa i}\mu_{\kappa i}) \right) \right],
\end{aligned}$$

where $j = 1,\ldots,p$. By equating these $2p+1$ equations to zero and then solving, we obtain the maximum likelihood estimates of the model parameters.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.

Arnold, J.M. and Straten, J.T. (2012). Motivation and skills as determinants of first-year performance in economics. *The Journal of Economic Education*, 43, 33–47.

Bayes, C.L., Bazán, J.L. and García, C. (2012). A new robust regression model for proportions. *Bayesian Analysis*, 4, 841–866.

Brooks, C. (2009). *RATS Handbook to Accompany Introductory Econometrics for Finance*. Cambridge University Press.

Cepeda-Cuervo, E. (2001). Modelagem de variabilidade em modelos lineares generalizados. *PhD thesis, Mathematics Institute, Federal University of Rio de Janeiro. Unpublished Thesis.*

Cepeda-Cuervo, C., Achcarb, J.A. and Garrido, L. (2014). Bivariate beta regression models: joint modeling of the mean, dispersion and association parameters. *Journal of Applied Statistics*, 41, 677–687.

Cepeda-Cuervo, E., Núñez-Antón, V. (2013). Spatial double generalized beta regression models: extensions and applications to study quality of education in Colombia. *Journal of Educational and Behavioral Statistics*, 38, 604–628.

Dolado, J.J. and Morales, E. (2009). Which factors determine academic performance of economics freshers? some spanish evidence. *Investigaciones Económicas*, 33, 179–210.

Ferrari, S. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31, 799–815.

Gómez-Déniz, E., Sordo, M. and Calderín-Ojeda, E. (2013). The log-Lindley distribution as an alternative to the beta regression model with applications in insurance *Insurance: Mathematics and Economics*, 54, 49–57.

Gottschalk, J. and Maslen, E. (1988). Reduction formulae for generalised hypergeometric functions of one variable. *Journal of Physics A: Mathematical and Theoretical*, 21, 1983–1998.

Gradshteyn, I. and Ryzhik, I. (1994). *Table of Integrals, Series, and Products*. Boston: Academic Press (5^{th} ed. A. Jeffrey).

Gurmu, S. and Elder, J. (2000). Generalized bivariate count data regression. models. *Economics Letters*, 68, 31–36.

Huang, X. and Oosterlee, C. (2011). Generalized beta regression models for random loss given default. *The Journal of Credit Risk*, 7, 1–26.

Jaynes, E. (1994). *Probability Theory: The Logic of Science*. `http://omega.math.albany.edu:8008/JaynesBook.html`.

Johnson, M. and Kuennen, E. (2006). Basic math skills and performance in an introductory statistics course. *Journal of Statistics Education*, 14, 20–40.

Lee, M.-L.T. (1996). Properties and applications of the Sarmanov family of bivariate distributions. *Communications in Statistics-Theory and Methods*, 25, 1207–1222.

Lehmann, E. (1996). Some concepts of dependence. *Annals of Mathematical Statistics*, 37, 1137–115.

Lunsford, M.L. and Poplin, P. (2011). From research to practice: basic math skills and success in introductory statistics. *Journal of Statistics Education*, 19, 1–22.

Maher, M. (1990). A bivariate negative binomial model to explain traffic accident migration. *Accident Analysis and Prevention*, 22, 487–498.

Olkin, I. and Liu, R. (2003). A bivariate beta distribution. *Statistics & Probability Letters*, 62, 407–412.

Paolino, P. (2001). Maximum likelihood estimation of models with beta distributed dependent variables. *Political Analysis*, 9, 325–346.

Papke, L. and Wooldridge, J. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics*, 11, 619–632.

Papke, L. and Wooldridge, J. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, 145, 121–133.

Pérez-Rodríguez, J.V. and Gómez-Déniz, E. (2015). Spread component costs and stock trading characteristics in the spanish stock exchange. *Quantitative Finance*, 15, 1943–1962.

Tong, Y. (1980). *Probability Inequalities in Multivariate Distributions*. Academic Press, New York.

Wolfram, S. (2003). *The Mathematica Book*. Wolfram Media, Inc.