

Implementation of Beta Regression Models on Concrete Strength Data with the Consideration of Variable Scale Parameters

Miaad Valipour Pashakolaei¹, Einolah Deiri^{2*}, Ezzatallah Baloui Jamkhaneh³

¹Ph.D. Student, Department of Statistics, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran

²PhD, Department of Statistics, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran *Corresponding author

³PhD, Department of Statistics, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran

Abstract

The nature of the variable in many regression usages is the response in terms of rate and ratio. As in economics, economists seek to understand the relationship between growth or unemployment rates and several other economic variables. Logistic and Probit models are usually used to model data with a variation range (0,1). However, Logistic and Probit models are not suitable for rate or ratio data modeling due to their concentration in a certain sub-range of their variation ranges. Regarding the high flexibility of the beta distribution for this type of data, a proper and efficient model is a regression based on beta distribution, which is called beta regression. This paper introduces the model and estimates its parameters. Then, we obtain an asymptotic distribution of estimators and evaluate the performance of the proposed model based on the MSE index using simulation. In the end, we will show the use of the model in a real example of concrete strength data.

Keywords: Concrete Strength Data, Beta Regression, Bond Function, Asymptotic Distribution, Generalized Linear Models

Introduction

An objective of statistical surveys is to find relationships that can be used to evaluate the effect of one or more variable changes on other variables (herman et al., 2019). Regression is the most common tool for achieving this goal. In many uses, the variable nature of the response is in terms of rate and ratio (Delcea & Siserman, 2020). In economics, for example, economists seek to understand the relationship between the rate of growth, unemployment, the percentage of national production, etc., and several other economic variables. In the industry, a manager seeks to find the factors that increase the defect rate of a device such as a turbine. In medicine, researchers seek to discover the proportion of the total lung cancer patients in a particular region (Voiță-Mekereș et al., 2023). In fact, the response variable in this data has a range at a distance of (0,1). Logistic and Probit models are usually used to model data with a variation range (0,1) (Lemeshow et al. 2013). The models are not suitable for their modeling since the rate or ratio data are usually concentrated in a certain sub-distance of their variation range (in other words, the distribution of this type of data is severely skewed). Thus, we provide a model known as the beta regression model with regard to great beta distribution flexibility. The main motive for introducing beta regression is the flexibility of beta distribution to consider the various types of distributions (0,1).

The statistical distribution of rates and ratios in the literature of statistical science has attracted many scholars and researchers (Murray et al. 2014). Various researchers have conducted studies based on beta regression models, the main idea of which was introduced by Williams (1982); however, the use of this model was welcomed after Ferrari and Cribari-Neto (2004). They carried out basic and comprehensive research on the beta

regression model based on the rate and ratio data and showed the usage of the beta regression model as well. They considered the accuracy parameter of the beta model to be constant and developed the maximum likelihood deduction. Simas et al. (2010) generalized the proposed model for a case where the parameter was not constant and was a function of observations. They analyze this kind of regression model from the Bayesian point of view. In this view, Branscum et al. (2007) also analyzed beta regression models. This regression model has found many advocates in recent years, including Paolini (2001), Smithson and Verkuilen (2006), Kelly et al. (2007), Espinheira (2008), Hahn (2008), Wallis et al. (2009), De Lima et al. (2016), Da Silva and de Oliveira Lima (2017), and Chen et al. (2017).

In this paper, we first introduce the beta regression model and estimate its parameters using the Maximum Likelihood method. In Section 3, we present an asymptotic distribution of estimators. In Section 4, we evaluate the asymptotic properties of the estimators using simulations. In Section 5, we implement the usage of the model on a real example that relates to concrete resistance data. In Section 6, we conclude and present the proposal (Figure 1).

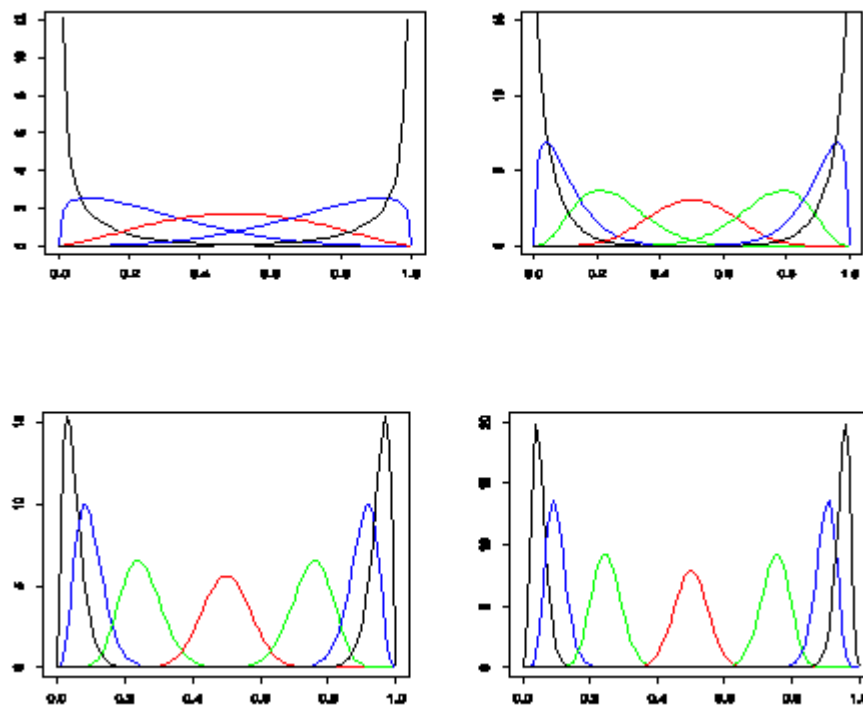


Figure 1. Beta density function graphs for different values of ϕ and μ .

1. Introduction of the beta regression model

Assume that the response variable y has a beta distribution with the parameters p and q . In this case, the density function y is:

$$f_{p,q}(y) = (\Gamma(p+q)/\Gamma(p)\Gamma(q))y^{p-1}(1-y)^{q-1}, \quad 0 < y < 1, \quad p > 0, \quad q > 0 \quad (1)$$

Typically, the regression model is formulated in terms of the mean of the response. A different form of the beta distribution density function is needed for an appropriate regression model based on the beta distribution, whose parameters are defined in terms of the mean μ and the accuracy parameter ϕ . For this purpose, if we set $\mu = p / (p + q)$ and $\phi = p + q$, we will have $q = (1-\mu)\phi$ and $p = \mu\phi$. Thus, it can be written as (Ferrari and Cribari-Neto 2004):

$$E(Y) = \mu \quad (2)$$

$$Var(Y) = \frac{\phi^2 \mu(1-\mu)}{\phi^2(\phi+1)} = \frac{\mu(1-\mu)}{\phi+1} = \frac{V(\mu)}{1+\phi} \quad (3)$$

In the expression $Var(Y)$, with the assumption of μ , the variance decreases if ϕ is increased. Therefore, ϕ can be interpreted as a model's accuracy estimation parameter. By defining new parameters, the density function of the response variable is as follows:

$$f(y|\mu, \phi) = (\Gamma(\phi)/\Gamma(\mu\phi)\Gamma((1-\mu)\phi))y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1, 0 < \mu < 1, \quad \phi > 0 \quad (4)$$

Wherein $\phi > 0$, $0 < \mu < 1$ and $0 < y < 1$. This type of beta display enables us to provide a regression model based on the category of generalized linear models (GLM). In fact, beta regression is a member of the GLM category. In this model, μ is a function of explanatory variables, and ϕ can be a constant or a function of variables (Simas et al. 2010). Suppose that the mean and accuracy parameters for μ_i are as follows:

$$g_1(\mu_i) = \eta_{1i} = x_i' \beta \quad g_2(\phi_i) = \eta_{2i} = z_i' \theta \quad (5)$$

in which $\beta = (\beta_1, \dots, \beta_k)'$ and $\theta = (\theta_1, \dots, \theta_h)'$ are vectors of unknown parameters for mean and accuracy functions, respectively, and $(\beta \in \mathbb{R}^k)$, $(\theta \in \mathbb{R}^h)$ and $k + h < n$, as well as the functions η_{1i} and η_{2i} are linear predictors, and $x_{i1}, \dots, x_{ik}$ and $z_{i1}, \dots, z_{ih}$ are the explanatory variables. Assume that the matrices X and Z are the matrices of the scheme corresponding to the explanatory variables $x_1, \dots, x_k$ and $z_1, \dots, z_h$.

2.1 Fitting the model

The logarithm of likelihood for a sample of n is expressed as follows for this beta regression model:

$$\ell(\beta, \phi) = \sum_{i=1}^n \ell_i(\mu_i, \phi) \quad (6)$$

$$\ell_i(\mu_i, \phi_i) = \log \Gamma(\phi_i) - \log \Gamma(\mu_i \phi_i) - \log \Gamma((1-\mu_i)\phi_i) + (\mu_i \phi_i - 1) \log y_i + \{(1-\mu_i)\phi_i - 1\} \log(1-y_i) \quad (7)$$

where the functions $g_1(\mu_i) = \eta_{1i}$ and $g_2(\phi_i) = \eta_{2i}$ defined in (6) are respectively the functions of β and θ . To estimate the parameters by the maximum likelihood method, we must obtain the corresponding function of β and ϕ . For the proof of theorem 1, a new approach is presented in the following.

Theorem 1: Consider the vector of the parameter $\xi = (\beta', \theta')$. We write the vector of the $(k + h)$ dimension of $U(\xi)$ as $(U_\beta(\beta, \theta)', U_\theta(\beta, \theta)')'$. The corresponding score functions of β and θ are:

$$U_\beta(\beta, \theta) = X' \phi T_1 (y^* - \mu^*) \quad (8)$$

$$U_\theta(\beta, \theta) = Z' T_2 v \quad (9)$$

where $T_1 = \text{diag}(\frac{d\mu_i}{d\eta_{1i}})$, $T_2 = \text{diag}(d\phi_i/d\eta_{2i})$, $\phi = \text{diag}(\phi_i)$, $y^* = (y_1^*, \dots, y_n^*)'$ 'And $\mu^* = (\mu_1^*, \dots, \mu_n^*)'$ such that $y_i^* = \log \frac{y_i}{1-y_i}$, and $\mu_i^* = \psi(\mu_i \phi) - \psi((1-\mu_i)\phi)$. The $\text{diag}(\mu_i)$, A n -dimensional matrix with μ_i components. We also have $v = (v_1, \dots, v_n)'$ so that

$$v_i = \mu_i(y_i^* - \mu_i^*) + \psi(\phi_i) - \psi((1-\mu_i)\phi_i) + \log(1-y_i).$$

Argument: By deriving from the logarithmic likelihood function, we have:

$$U_r(\beta, \theta) = \frac{\partial \ell(\beta, \theta)}{\partial \beta_r} = \sum_{i=1}^n \frac{\partial \ell_i(\mu_i, \phi_i)}{\partial \mu_i} \left[\left(\frac{d\mu_i}{d\eta_{1i}} \right) \left(\frac{\partial \eta_{1i}}{\partial \beta_r} \right) \right] \quad (10)$$

According to the definition of $(d/dz)\log \Gamma(z) = \psi(z)$:

$$\partial \ell_i(\mu_i, \phi_i) / \partial \mu_i = [\log \frac{y_i}{1-y_i} - \{\psi(\mu_i \phi_i) + \psi((1-\mu_i)\phi_i)\}] \quad (11)$$

By replacing the equations, we get:

$$U_r(\beta, \phi) = \sum_{i=1}^n \phi_i ((y_i^* - \mu_i^*) / g'_1(\mu_i)) x_{ir} \quad (12)$$

Therefore:

$$U_\beta = \begin{pmatrix} U_1 \\ \vdots \\ U_k \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n \phi_i \frac{(y_i^* - \mu_i^*)}{g'_1(\mu_i)} x_{i1} \\ \vdots \\ \sum_{i=1}^n \phi_i \frac{(y_i^* - \mu_i^*)}{g'_1(\mu_i)} x_{ik} \end{pmatrix} = X' \phi T_1 (y^* - \mu^*).$$

Similarly, by calculating the θ_R rating function, we have:

$$U_\theta = \begin{pmatrix} U_1 \\ \vdots \\ U_h \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n \frac{v_i}{g'(\phi_i)} z_{i1} \\ \vdots \\ \sum_{i=1}^n \frac{v_i}{g'(\phi_i)} z_{ih} \end{pmatrix} = Z' T_1 v.$$

To complete the model's fitting process, we need to know the accuracy of the obtained estimators. Thus, we require calculating Fisher's knowledge matrix of estimators. The inverse of this matrix results in a covariance matrix of model estimators.

Theorem 2: The Fisher Informational Matrix for the model parameters (4) is $K(\xi) = P'WP$, in which:

$$P = \begin{bmatrix} X & 0 \\ 0 & Z \end{bmatrix}$$

and W is a $2n \times 2n$ matrix as follows:

$$W = \begin{bmatrix} W_{\beta\beta} & W_{\beta\theta} \\ W_{\beta\theta} & W_{\theta\theta} \end{bmatrix}$$

so that

$$W_{\beta\beta} = \text{diag} \left(\phi_i^2 a_i \left(\frac{1}{g'_2(\mu_i)} \right)^2 \right),$$

$$W_{\beta\theta} = \text{diag} \left(\phi_i \left\{ \mu_i c_i - \frac{d}{d\phi_i} \psi((1-\mu_i)\phi_i) \right\} \left(\frac{1}{g'_1(\mu_i)} \right) \left(\frac{1}{g'_2(\phi_i)} \right) \right),$$

$$W_{\theta\theta} = \text{diag} \left(b_i \left(\frac{1}{g'_2(\phi_i)} \right)^2 \right),$$

where

$$a_i = \frac{d}{d\mu_i} \psi((1-\mu_i)\phi) + \frac{d}{d\mu_i} \psi(\mu_i \phi_i),$$

$$b_i = \frac{d}{d\phi_i} \psi((1 - \mu_i)\phi)(1 - \mu_i)^2 + \frac{d}{d\phi_i} \psi(\mu_i\phi_i)\mu_i^2 - \frac{d}{d\phi_i} \psi(\phi_i),$$

$$c_i = \frac{d}{d\phi_i} \psi((1 - \mu_i)\phi) + \frac{d}{d\phi_i} \psi(\mu_i\phi_i).$$

Argument: By deriving Equation (10) by β , we have:

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \theta)}{\partial \beta_r \partial \beta_s} &= \sum_{i=1}^n \frac{\partial}{\partial \mu_i} \left(\frac{\partial \ell_i(\mu_i, \phi_i)}{\partial \mu_i} \frac{d\mu_i}{d\eta_{1i}} \right) \frac{d\mu_i}{d\eta_{1i}} \frac{\partial \eta_{1i}}{\partial \beta_s} x_{ir} \\ &= \sum_{i=1}^n \left(\frac{\partial^2 \ell_i(\beta, \theta)}{\partial \mu_i^2} \frac{d\mu_i}{d\eta_{1i}} + \frac{\partial \ell_i(\beta, \theta)}{\partial \mu_i} \frac{\partial}{\partial \mu_i} \frac{d\mu_i}{d\eta_{1i}} \right) \frac{d\mu_i}{d\eta_{1i}} x_{ir} x_{is}. \end{aligned}$$

Since $E\left(\frac{\partial \ell_i(\mu_i, \phi_i)}{\partial \mu_i}\right) = 0$, then

$$E\left(\frac{\partial^2 \ell(\beta, \theta)}{\partial \beta_r \partial \beta_s}\right) = \sum_{i=1}^n \left(\frac{\partial^2 \ell_i(\mu_i, \theta_i)}{\partial \mu_i^2} \right) \left(\frac{d\mu_i}{d\eta_{1i}} \right)^2 x_{ir} x_{is}.$$

Deriving from the derivation of (11) results in:

$$\begin{aligned} \frac{\partial^2 \ell_i(\mu_i, \phi_i)}{\partial \mu_i^2} &= \frac{d}{d\eta_{1i}} \left(\phi_i \left[\log \frac{y_i}{1 - y_i} - \{\psi(\mu_i\phi_i) - \psi((1 - \mu_i)\phi_i)\} \right] \right) \\ &= \phi_i \left(-\phi_i \frac{d}{d\mu_i} \psi(\mu_i\phi_i) - \phi_i \frac{d}{d\mu_i} \psi((1 - \mu_i)\phi_i) \right) \\ &= \phi_i^2 \left(\frac{d}{d\mu_i} \psi(\mu_i\phi_i) - \frac{d}{d\mu_i} \psi((1 - \mu_i)\phi_i) \right). \end{aligned}$$

According to the definition a_i , we have:

$$E\left(\frac{\partial^2 \ell(\beta, \theta)}{\partial \beta_r \partial \beta_s}\right) = - \sum_{i=1}^n (\phi_i^2 a_i) \left(\frac{1}{g'_1(\mu_i)} \right)^2 x_{ir} x_{is}.$$

As a result:

$$-E\left(\frac{\partial^2 \ell(\beta, \theta)}{\partial \beta \partial \beta'}\right) = X' \text{diag} \left(\phi_i^2 a_i \left(\frac{1}{g'_1(\mu_i)} \right)^2 \right) X = X' W_{\beta\beta} X.$$

Other components of the Fisher Informational Matrix are similar to those above.

Maximum Likelihood Estimates of Beta Regression Parameters are not analytically computable and should be obtained by numerical methods. These estimators are calculated by the betareg package in R software (Cribari-Neto and Zeileis 2010).

2. Asymptotic distribution of estimators

In this section, we examine the asymptotic characteristics of the estimators and introduce confidence intervals for them. The theorem below shows the asymptotic distribution of estimators.

Theorem 3: (Casella and Berger 2002) When the sample size is large, and assuming the existence and reversible of $J(\xi) = \lim_{n \rightarrow \infty} \frac{K(\xi)}{n}$, under the terms of the order, the asymptotic distribution of the maximum likelihood estimators of the vector parameter ξ is:

$$\sqrt{n}(\hat{\xi} - \xi) \sim N_{k+h}(0, J(\xi)^{-1}).$$

Using this, the asymptotic confidence intervals of estimators can be calculated as follows. If $K(\hat{\xi})^{rr}$, the r^{th} of component of the matrix $K(\hat{\xi})^{-1}$ and $K(\hat{\xi})^{RR}$, $(K + R)$, the diameter of the component $K(\hat{\xi})^{-1}$, and then the asymptotic confidence intervals of 100 $(1 - \alpha)\%$ for β_r and θ_R are obtained as follows.

$$\widehat{\beta}_r \mp \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) (K(\hat{\xi}^{rr})^{\frac{1}{2}}),$$

$$\widehat{\theta}_R \mp \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) (K(\hat{\xi}^{rr})^{\frac{1}{2}}),$$

3. Simulation study

With a simulation study in the small sample, we will examine the compatibility and normal asymmetry of model estimators for a case where ϕ is a function of explanatory variables. We consider the model as follows:

$$\eta_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2, \quad \eta_2 = \beta_0 + \theta_1 z_1 + \theta_2 z_2, \quad i = 1, \dots, n, \quad (13)$$

in which the values of the explanatory variables x_1 and x_2 were respectively generated from distributions $\mathcal{N}(0,1)$ and Bernoulli with a probability of success of 0.6. The values of the explanatory variables z_1 and z_2 were also generated respectively from Bernoulli distributions with a probability of success of 0.5 and $\mathcal{N}(1,2)$. We produced a collection of data from model (12) with real values $(\beta_0, \beta_1, \beta_2) = (-1, 0.75, 1.5)$ and $(\theta_0, \theta_1, \theta_2) = (2, -1, 0.5)$, and we obtained the maximum likelihood estimates for the parameters of each. Using the vector of estimated parameters, we calculated the error values and Mean Squared Error (MSE). The results are reported in Table 1.

Table 1. Simulation results for estimating coefficients of beta regression with the logit bonding function for mean and logarithms for the accuracy parameter using 500 data sets.

Sample Size	Parameters	Real Quantity	Estimation	MSE	Bias
30	β_0	-1	-1.0188	-0.0407	0.0407
	β_1	0.75	0.7558	-0.0209	0.0209
	β_2	1.5	1.5331	-0.0331	0.0591
	θ_0	2	2.2430	0.2430	0.3042
	θ_1	-1	-1.1278	-0.1278	0.4496
	θ_2	0.5	0.5569	0.0569	0.0262
50	β_0	-1	-0.9981	-0.0018	0.0218
	β_1	0.75	0.7490	-0.0009	0.0103
	β_2	1.5	1.4983	-0.0016	0.0349
	θ_0	2	2.1159	0.1159	0.1137
	θ_1	-1	-1.0404	-0.0404	0.1714
	θ_2	0.5	0.5262	0.0262	0.0146
100	β_0	-1	-1.0026	-0.0026	0.0090
	β_1	0.75	0.7515	-0.0015	0.0047
	β_2	1.5	1.4951	-0.0048	0.0153

	θ_0	2	2.0575	0.0575	0.0445
	θ_1	-1	-1.0156	0.0156-	0.0715
	θ_2	0.5	0.5119	0.0119	0.0052
200	β_0	-1	-1.0033	-0.0033	0.0044
	β_1	0.75	0.7512	-0.0012	0.0017
	β_2	1.5	1.5040	-0.0040	0.0072
	θ_0	2	2.0311	0.0311	0.0226
	θ_1	-1	-1.0144	-0.0144	0.0360
	θ_2	0.5	0.5056	-0.0056	0.0022
400	β_0	-1	-0.9988	-0.0015	0.0022
	β_1	0.75	0.7497	-0.0015	0.0009
	β_2	1.5	1.4996	-0.0009	0.0038
	θ_0	2	2.0126	0.0218	0.0105
	θ_1	-1	-0.9997	-0.0098	0.0171
	θ_2	0.5	0.0504	-0.0025	0.0010

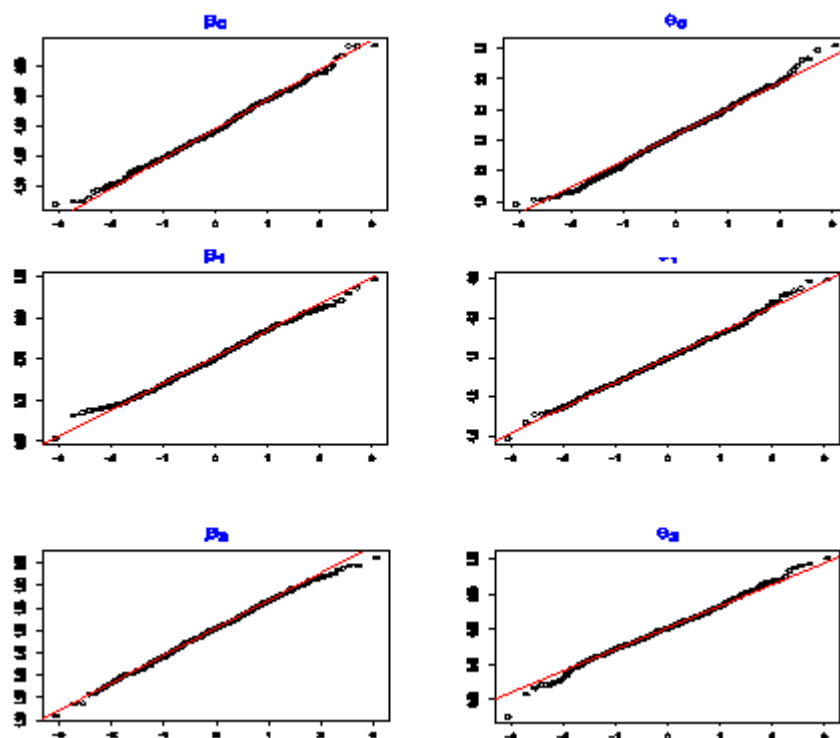


Figure 2. Quantile-quantile plots for the model coefficients (13).

According to the results of Table 1, the MSE estimates decrease with the increase in sample size. On the other hand, the bias tends to be zero. The odd values appear to be unbiased. Consequently, the compatibility of estimators can be deduced. According to Theorem 3, the asymptotic distribution of estimators is normal. We use multi-quintile graphs (Figure 2) to examine this feature in a small sample. The asymptomatic estimators are verified to be normal according to Figure 2.

4. Formation of concrete strength data

In this section, we fit the introduced beta regression model onto concrete resistance data. In this data, the response variable is the percentage of hardness (strength) of the concrete. The purpose of the data analysis is to investigate the effect of micro silica, sand, cement, and water as the model variables on concrete hardness. Rezaei (2012) reported this dataset.

To accurately determine the model, we first sought explanatory variables using an exploratory analysis that could be present in the linear predictor of the μ and ϕ parameters. We use all variables for the μ parameter. To determine the possible candidates for attending ϕ , we use the box plot of the response variable in the categories of explanatory variables. Figure 3 shows these charts for all explanatory variables. According to this figure, the response variables are different in the categories of sand, cement, and water variables. Nevertheless, these changes remain almost the same in the different categories of microsilica variables. Therefore, the linear predicate for ϕ is the explanatory variables of sand, water, and cement as primary candidates. In summary, the beta regression model for these data is as follows:

$$\text{logit}(\mu_i) = x_i'\beta, \quad \log(\phi_i) = z_i'\theta \quad i = 1, \dots, 108,$$

Where

$$x_i = (1, \text{sand}, \text{water}, \text{Micro silica}, \text{Cement}), \quad z_i = (1, \text{Sand}, \text{Water}, \text{Cement})$$

$$\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4), \quad \theta = (\theta_0, \theta_1, \theta_2, \theta_3)$$

Table 2. Estimated Parameters Using Concrete Construction Data with the Variable ϕ .

Model	Coefficients	Parameter	Estimation	Wald Statistic	Standard error	p-value
μ	Width from origin	β_0	-11.8	-9.98	1.110	<0.0001***
	Sand	β_1	0.01	9.24	0.008	<0.0001***
	Water	β_2	-0.01	-5.26	0.001	<0.0001***
	Micro silica	β_3	0.02	9.61	0.001	<0.0001***
	Cement	β_4	0.01	13.44	0.009	<0.0001***
ϕ	Width from origin	θ_0	-10.89	-3.42	3.170	<0.0001***
	Sand	θ_1	0.01	2.34	0.002	<0.0001***
	Water	θ_2	0.02	2.76	0.005	0.0059**
	Cement	θ_3	0.02	5.27	0.003	<0.0001***

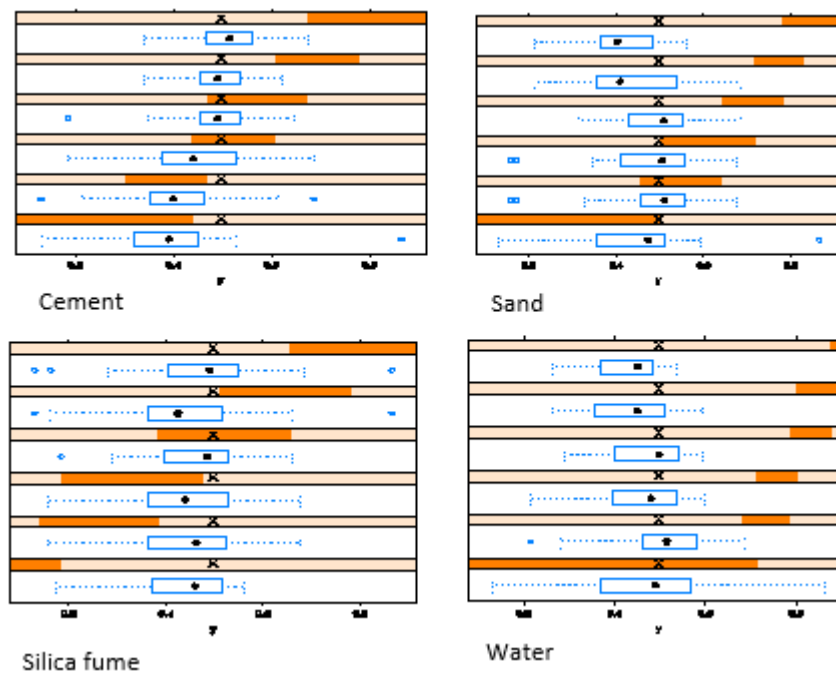


Figure 3. Box variables for response variables in different categories of explanatory variables for concrete strength data.

5. Discussion and conclusion

According to the fitting results of the model in Table 2, all coefficients in the model are significant considering the obtained p-value, which is less than the contract value of 0.00. Moreover, only the water explanatory variable has a reverse effect on the concrete hardness, and an increase in the amount of water decreases concrete strength. Similarly, the greatest effect on the average concrete hardness was due to the microsilica variable. In other words, the amount of microsilica in concrete strength significantly affects the other concrete ingredients. Finally, the degree of model accuracy increases with increasing levels of variables for all the explanatory variables in the linear predicate of ϕ .

In this paper, an appropriate regression model is presented for responses that are continuously changed at the distance of (0,1). This kind of data includes the rates and ratios that researchers encounter in different usage situations. In the introduced models, the basic acceptance is that the response variable follows the beta exponential distribution. It is well known that the beta distribution is very flexible for modeling data at unit intervals because the beta density function shows different curves with different shape parameters.

The simulation study shows that the asymptomatic behavior of estimators is consistent with the theoretical results, which are available even for small and medium-sized volumes.

References

- [1] Atkinson, A. C. 1985. *Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*. New York: Oxford University Press.
- [2] Bickel, P. J., and K. J. Doksum. 2001. *Mathematical Statistics*. Prentice-Hall New Jersey.
- [3] Branscum, A. J., W. O. Johnson, and M. C. Thurnond. 2007. "Bayesian beta regression application to household data genetic distance between foot-and-mouth disease viruses." *Australian and New Zealand Journal of Statistics* 49 (3): 287-301.
- [4] Chen, K. K., R. H. Chiu, and Ch. T. Chang. 2017. "Using beta regression to explore the relationship between service attributes and likelihood of customer retention for the container shipping industry." *Transportation Research Part E* 104: 1-16.

- [5] Cribari-Neto, F., and A. Zeilis. 2010. "Beta regression in R." *Journal of Statistical Software* 34: 54-71.
- [6] Da Silva, A. R., and A. De Oliveira Lima. 2017. "Geographically Weighted Beta Regression." *Spatial Statistics* 21: 279-303.
- [7] De Lima, M. S., V. S. Dos Santos, L. H. Duczmal, and D. D. S. Souza. 2016. "A spatial scan statistic for beta regression." *Spatial Statistics* 18: 444-454.
- [8] Delcea, C., & Siserman, C. (2020). Validation and standardization of the questionnaire for evaluation of paraphilic disorders. *Romanian Journal of Legal Medicine*, 28(1), 14–20. <https://doi.org/10.4323/rjlm.2020.14>
- [9] Espinheira, P., S. Ferrari, and F. Cribari-Neto. 2008. Influence diagnostics in beta regression. *Computational Statistics and Data Analysis* 52 (9): 4417-4431.
- [10] Ferrari, S. L. P., and F. Cribari-Neto. 2004. "Beta Regression for Modeling Rates and Proportion." *Journal of Applied Statistics* 31: 799-815.
- [11] Hahn, E. D. 2008. "Mixtrue densities for project management activity times: A robust approach to PERT." *European Journal of Operational Research* 188: 450-459.
- [12] herman, C., Enache, A., & Delcea, C. (2019). The multi-factorial determinism of forensic expertise regarding sentence interruption on medical grounds and decision. *Journal of Forensic and Legal Medicine*, 61, 45–55. <https://doi.org/10.1016/j.jflm.2018.11.010>
- [13] Kelly, G., R. Garabed, A. Branscum, A. Perez, and M. Thurmomd. 2007. "Prediction model for sequence variation in the glycoprotein gene of infectious hematopoietic necrosis virus in California, USA." *Diseases of Aquatic Organisms* 78: 97-104.
- [14] Koenker, R. 2005. *Quantile Regression*. Cambridge University Press.
- [15] Lemeshow, S., D. W. Hosmer Jr., and R. X. Sturdivant. 2013. *Applied Logistic Regression*. John Wiley and Sons.
- [16] Murray, A. T., T. H. Grubestic, and R. Wei. 2014. "Spatially significant cluster detection." *Spatial Statistics* 10: 103-116.
- [17] Paolini, P. 2001. "Maximum likelihood estimation of models with beta-distribution dependent variable." *Political Analysis* 9: 325-346.
- [18] Rezaei M. 2012. "Estimation of the fundamental properties of hsc concrete using statistical methods and artificial neural networks." MSc thesis, Sahand Industrial University.
- [19] Simas, A. B., W. Barreto-Souza, and A. V. Rocha. 2010. "Improved estimators for a general class of beta regression models." *Computational Statistics and data Analysis* 54:348-366.
- [20] Smithson, M., and J. Verkuilen. 2006. "A better Lemon Squeezer? Maximum likelihood Regression with Beta-Distribution Dependent Variables." *Psychological Methods* 11 (1): 54-71.
- [21] Voiță-Mekereș, F., Delcea, C., Buhaș, C. L., & Ciocan, V. (2023). Novichok toxicology: A review study. *Archives of Pharmacy Practice*, 14(3), 62–66. <https://doi.org/10.46542/APP.2023.142>
- [22] Wald, A. 1943. "Test of Statistical Hypotheses Concerning Several Parameters When the Number of Obseervations is Large." *Transcations of the American Mathematical Society* 54: 426–482.
- [23] Wallis, E., R. Nally, and S. Lake. 2009. "Do tributaries affect loads and fluxes of particulate organic matter, inorganic sediment and wood? Patterns in an upland river basin in south-eastern Australia." *Hydrobiologia* 636: 307-317.
- [24] Williams, D. A. 1982. "Extra binomial variation in logistic linear models." *Applied Statistics* 31: 144-148.
- [25] Yanqing, Y., and X. Wang. 2011. "Comparsion of Wald, Score, and Likelihood Ratio Tests for Response Adaptive Designs." *Journal of Statistial Theory and Applications* 10: 553-569.