Afro-Asian Journal of Scientific Research (AAJSR)
المجلة الأفروآسيوية للبحث العلمي
E-ISSN: 2959-6505
Volume 4, Issue 1, 2026
Page No: 82-88
Website: https://aajsr.com/index.php/aajsr/index
SJIFactor 2024: 5.028 ISI 2025: 0.915 معامل التأثير العربي (AIF) 2025: 0.76

Enhancing Beta Regression for Bounded Response Modeling Using Spline-Based Mean and Precision Functions

Abobaker M. Jaber[1*], Mariam.A. Orafi[2], Ahmed M. Mami[3], Naeima N. Abd Elati، Salma Saad[5]
[1,2,4,5]Statistics Department, University of Benghazi, Benghazi, Libya
[3]Mathematics Department, University of Omar Al-Mukhtar, Al Bayda, Libya

تحسين نماذج انحدار بيتا للمتغيرات المحدودة القيمة باستخدام دوال انسيابية للمتوسط والدقة

أبوبكر محمد جابر[1*]، مريم عبدالله العرفي[2]، أحمد محمد مامي[3]، نعيمة نصر عبد العاطي[4]، سالمة سعد[5]
[1،2،4،5]قسم الاحصاء، كلية العلوم، جامعة بنغازي، بنغازي، ليبيا
[3]قسم الرياضيات، كلية العلوم، جامعة عمر المختار، البيضاء، ليبيا

*Corresponding author: jaber@uob.edu.ly

Received: December 07, 2025	Accepted: January 19, 2026	Published: January 28, 2026

Abstract:
Beta regression is widely used for modeling continuous outcomes constrained to the unit interval (0,1). Classical beta regression often assumes linear relationships between predictors and the mean or precision parameter, potentially limiting flexibility for complex data patterns. We propose a spline-enhanced beta regression framework that models both the mean and precision parameters using natural cubic splines. We benchmark this method against Transformed Gaussian regression, Quantile regression, using three diverse test functions: Trigonometric, Polynomial, and Exponential-Log. Extensive simulations over multiple sample sizes demonstrate that the spline-based beta regression consistently improves predictive accuracy while maintaining model interpretability. Our results suggest that spline-based extensions provide a robust and flexible alternative for bounded response modeling.

Keywords: Beta regression, Spline modeling, Bounded outcomes, Predictive accuracy, Natural cubic splines.

الملخص:
يُعدّ انحدار بيتا أحد الأساليب الإحصائية الشائعة لنمذجة المتغيرات المستمرة المحدودة (المحصورة) ضمن المجال (0,1). إلا أنّ النماذج التقليدية غالباً ما تفترض وجود علاقات خطية بين المتغيرات التفسيرية وكلٍّ من متوسط التوزيع أو معامل الدقة، وهو ما قد يقيّد قدرتها على تمثيل الأنماط البيانية المعقدة. وفي هذا العمل، نقترح نموذجاً مطوّراً لانحدار بيتا يعتمد على دوال انسيابية تكعيبية للمتوسط والدقة لنمذجة كلٍّ من متوسط الاستجابة ومعامل الدقة في آن واحد .وقد قمنا بمقارنة أداء النموذج المقترح مع كل من الانحدار الغاوسي المحوَّل والانحدار الكمي، مستخدمين ثلاث دوال اختبار متنوعة :دالة مثلثية، ودالة متعددة الحدود، ودالة أسّية-لوغاريتمية. وتُظهر نتائج المحاكاة الواسعة عبر أحجام عينات متعددة أن نموذج انحدار بيتا المعزَّز بالسبلاين يوفر تحسناً ملحوظاً في دقة التنبؤ، مع المحافظة على قابلية تفسير النموذج. وتشير النتائج إلى أنّ نماذج الانحدار المعتمدة على السبلاين تمثّل بديلاً مرناً وقوياً لنمذجة المتغيرات المحصورة ضمن مجال محدود.

الكلمات المفتاحية: انحدار بيتا، النمذجة الإحصائية، المتغيرات المحدودة، الدوال الانسيابية، معامل الدقة.

Introduction:

Many scientific disciplines require the modeling of proportions, rates, or percentages bounded between 0 and 1. Examples include success rates in clinical trials, relative abundances in ecology, and normalized performance metrics in engineering. Traditional linear regression is unsuitable for such outcomes due to the bounded nature of the response variable, which violates assumptions of normality and homoscedasticity. Beta regression, introduced by Ferrari and Cribari-Neto (2004), provides a principled alternative by modeling outcomes $(Y \in (0,1))$ using the beta distribution, parameterized by a mean $(\mu)$and a precision $(\emptyset)$. Despite its widespread adoption, standard beta regression typically assumes linear or log-linear relationships between covariates and both parameters. This structural limitation may fail to capture nonlinear dependencies often present in complex, real-world datasets, thereby restricting model flexibility and predictive accuracy. In response, this study proposes a spline-enhanced beta regression framework that extends the classical approach by allowing flexible, nonlinear modeling of both the mean and precision parameters using natural cubic splines. Our method preserves the interpretability of regression models while accommodating intricate functional forms without prior specification. The remainder of this paper is organized as follows. Section 2 reviews related work on beta regression and flexible modeling approaches. Section 3 details the proposed methodology, including classical beta regression, spline-based extensions, and competing methods. Section 4 describes the simulation design used to evaluate performance. Section 5 presents numerical and visual results, and Section 6 discusses their implications. Finally, Section 7 concludes with a summary and directions for future research.

Related Work:

Beta regression has been extensively studied for bounded data (Ferrari & Cribari-Neto, 2004). Extensions include modeling precision as a function of covariates (Smithson & Verkuilen, 2006) and Bayesian formulations (Ma et al., 2019). Spline-based regression techniques, particularly natural cubic splines, have been shown to capture non-linear patterns in generalized linear models (Wood, 2017). Recent studies have explored neural network extensions for bounded outputs (Wang et al., 2020), but such methods often lack interpretability.

Methodology:

Classical Beta Regression:

The beta distribution for (Y∈ (0,1)) is:

$$\left[f\,(y;\,\mu,\emptyset)=\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma\left((1\text{-}\mu)\phi\right)}\,y^{\mu\,\phi\text{-}1}(1\text{-}y)^{(1\text{-}y)\phi\text{-}1}\right]$$

where (μ ∈ (0,1)) is the mean, (ϕ > 0) is the precision. In classical beta regression:

$$\left[g\,(\mu_i)=X_i^{\top}\beta,\ \ \mathrm{h}\,(\emptyset_i)=Z_i^{\top}\gamma\right]$$

with link functions (g) and (h).

Spline-Based Beta Regression:

We propose replacing linear predictors with **natural cubic splines**:

$$\left[g\,(\mu_i)=f_\mu\,(x_{i1})+f_\mu\,(x_{i2}),\qquad \mathrm{h}(\phi_i)=f_\phi(x_{i1})+f_\phi(x_{i2})\right]$$

where (f_μ) and (f_ϕ) are spline functions.

Proposed Method: Spline-Based Beta Regression – Step-by-Step

Step 1: Model the Response Distribution

Assume the response variable $(Y \in (0,1))$ follows a beta distribution parameterized by mean $(\mu)$and precision $(\emptyset)$:

$$\left[Y{\sim}Beta\big(\mu\emptyset,(1-\mu)\big)\emptyset\right]$$

where:

- ($\mu \in (0,1)$) is the mean parameter.
- ($\emptyset > 0$) is the precision parameter.

Step 2: Replace Linear Predictors with Splines

In classical beta regression, linear predictors are used for both mean and precision:

$$[g(\mu_i) = X_i^{\top}\beta,\quad h(\phi_i) = Z_i^{\top}\gamma]$$

In the proposed method, these are replaced with natural cubic splines:

$$g(\mu_i) = f_\mu(x_{i1}) + f_\mu(x_{i2}) + \ldots$$
$$h(\phi) = f_\phi(x_{i1}) + f_\phi(x_{i2}) + \ldots$$

where:

- (f_μ) and (f_ϕ) are smooth spline functions,
- ($x_{i1}, x_{i2}, \ldots$) are covariates,
- $g(.)$ and $h(.)$ are appropriate link functions (e.g., logit for (μ_i), log for (ϕ).

Step 3: Implement Using Natural Cubic Splines
- Use natural cubic spline bases (e.g., via the `ns()` function in R) to represent (f_μ) and (f_ϕ).
- The splines allow flexible, nonlinear relationships between covariates and both the mean and precision parameters.
Step 4: Model Fitting and Estimation
- Estimate the spline coefficients using maximum likelihood estimation (MLE) or Bayesian methods.
- The likelihood function is based on the beta density:

$$\left[f\ (y\ ; \mu,\emptyset)=\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma\left((1-\mu)\phi\right)} y^{\ \mu\,\phi-1}(1-y)^{\ (1-y)\phi-1}\right]$$

- Optimization can be performed using standard statistical software (e.g., `betareg` with spline terms in R).

Competing Methods:

Transformed Gaussian Regression Models:

A traditional econometric approach for fractional data, as discussed by Papke and Wooldridge (1996), involves applying a transformation $h(.)$ (e.g., logit, probit) to map the bounded response to the real line and then employing a linear Gaussian model:

$$h(y_i)=X_i^T\beta+\epsilon_i, \qquad \epsilon\sim N(0,\, \sigma^2)$$

Papke and Wooldridge critically highlighted the fundamental flaws of this approach. First, the model is linear on the transformed scale, not the original scale of the data. Back-transforming the predictions yields an estimate of the conditional median of y, not the conditional mean, due to the nonlinear transformation. Recovering the mean requires retransformation that depends on the often-unknowable distribution of the errors, leading to biased estimates. Second, this method implicitly assumes constant variance on the transformed scale, which rarely holds, leading to inefficient estimates and invalid inference.

Quantile Regression Models for Bounded Data:

Quantile regression, as applied to bounded data in works like Taylor (2019), offers a non-parametric alternative that models the conditional quantiles of the response. For a quantile τ the model is:

$$\left[Q_{y_i}(\tau|X_i)=X_i^T\beta(\tau)\right]$$

where $Q_{y_i}(\tau|X_i)$ is the τ conditional quantile. This approach is highly robust and makes no distributional assumptions about the response. Its primary strength lies in characterizing the entire conditional distribution, which is particularly valuable for understanding tail behavior in bounded data. However, a key limitation is that quantile regression does not provide a full probabilistic model; it models quantiles individually, which can lead to "crossing" quantiles if not constrained. Furthermore, for a full distributional characterization, multiple quantiles must be estimated, which can be computationally intensive.

Simulation Design:

To evaluate how different modelling approaches perform when estimating mean functions for Beta-distributed outcomes, we designed a simulation study using three distinct data-generating mechanisms. Each mechanism specifies a nonlinear mean surface, ($\mu(x_1,x_2)$, and a corresponding precision surface, ($\emptyset(x_1,x_2)$.

1. Trigonometric: $\left(\boldsymbol{\mu=logit^{-1}\left(sin\ (2\pi x_1)+cos\ (2\pi x_2)\right),\ \ \phi_i=10+5sin\ (\pi x_1 x_2)}\ \right)$
2. Polynomial: $\left(\boldsymbol{\mu=logit^{-1}\left(2x_1-2x_2+4x_1x_2-2x_1^2+1.5x_2^2\right),\ \ \phi_i=8+3x_1-2x_2+6x_1x_2}\ \right)$
3. Exponential-Log: $\left(\boldsymbol{\mu=logit^{-1}(exp(0.5x_1)+log(1+2x_2)-1),\ \phi_i=12+4e^{-2x_1}+3log\ (1+3x_2)}\right)$

were selected to represent qualitatively different forms of nonlinearity: oscillatory behaviour, polynomial curvature, and exponential–logarithmic structure. These choices allow us to test whether the regression methods can adapt to a wide range of underlying patterns. For each mechanism, we generated samples of size (n) {250, 500, 750}. Covariates (x_1) and (x_2) were drawn independently from a uniform distribution on ([0,1]). We then computed the true mean $\boldsymbol{\mu(x_1,\ x_2)}$ and precision $\boldsymbol{\phi(x_1,\ x_2)}$ values and simulated responses ($Y\sim Beta(\mu\phi,\ (1-\mu)\phi)$). The function simulate data in the R code carries out this process and returns a dataset containing the observed outcomes, covariates, and the corresponding true generative values. To assess predictive accuracy, each dataset was analyzed using four different modelling strategies under 5-fold cross-validation:

1. **Beta Regression with Splines**, allowing flexible nonlinear effects in both mean and precision;
2. **Transformed Gaussian Regression**, fitting a linear model to the logit-transformed response;
3. **Median Quantile Regression** at (τ = 0.5);
4. **Classical Beta Regression** with linear predictors.

Within each fold, the models were trained on 80% of the data and evaluated on the remaining 20%. For each method, we extracted predictions of the conditional mean and compared them with the true generative values. Performance was summarized using **Root Mean Squared Error (RMSE)** and **Mean Absolute Error (MAE)**. After 500 runs were completed, the results were aggregated to obtain mean and standard deviation values for RMSE and MAE.

Results:

Discussion of Numerical Results:

Table 1. illustrate Mean Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) obtained from 500 run for each combination of test function, sample size, and modelling approach. Results are reported for the Beta-Spline regression, Classical Beta regression, Quantile Regression (τ=0.5), and Transformed Gaussian models. Lower values indicate better predictive performance.

Comparison Across Effects of Sample Size:

Across all generating mechanisms and modelling strategies, increasing the sample size from (n=250), (n=500), and (n=750) consistently reduces both RMSE and MAE. This behaviour is expected, as larger samples provide more information about the underlying mean surface, allowing the fitted models to better approximate the true data-generating process. The improvement is most pronounced for models with flexible components,such as the Beta-Spline model, which benefit substantially from additional data to stabilize spline-based estimates. More rigid models (e.g., classical beta regression) show smaller improvements, reflecting their limited ability to adapt to complex nonlinear structures even when sample size increases.

Comparison Across Test Functions:

1. Exponential–Log Mechanism:

All models perform well under this generating mechanism, with RMSE values generally between 0.011 and 0.024 across sample sizes. The data-generating mean surface is smooth and moderately nonlinear, making it easier for all methods to approximate. At every sample size, the Classical Beta Regression yields the best performance with the lowest RMSE and MAE. For example, at (n=750), it attains an RMSE of 0.0117 and an MAE of 0.0091, slightly outperforming the Beta-Spline model.

2. Polynomial Mechanism:

Errors are substantially larger in this setting, with RMSE values between approximately 0.072 and 0.083. This reflects the high curvature of the polynomial mean surface. Here, the **Beta-Spline** and **Quantile Regression** models consistently outperform the Classical Beta Regression and the Transformed Gaussian model. For example, at (n=750), the best RMSE (0.0723) is achieved by the Beta-Spline model, closely followed by Quantile Regression (0.0727). The Classical Beta Regression performs worst, with RMSE values exceeding 0.081 across all sample sizes.

3. Trigonometric Mechanism:

This is the most challenging scenario due to its oscillatory structure. The **Beta-Spline model** is the only method capable of adequately capturing this nonlinearity. Its RMSE values range from 0.0415 (at (n=250)) down to 0.0359 (at (n=750)), far outperforming all other methods. The Classical Beta Regression performs extremely poorly, with RMSE values consistently around 0.178, indicating its inability to adapt to highly periodic structures. Quantile Regression and the Transformed Gaussian approach perform moderately, but remain far inferior to the spline-based method.

5.1.3 Best-Performing Models (RMSE and MAE)

- **Exponential–Log Function:**
  Best Model: **Classical Beta Regression** *Best Performance:* RMSE = 0.0117, MAE = 0.0091 (at (n=750)) The linear structure embedded in the generating process aligns well with the assumptions of classical beta regression.
- **Polynomial Function:**
  Best Model: **Beta-Spline Regression** (closely followed by Quantile Regression) *Best Performance:* RMSE = 0.0723, MAE = 0.0552 (at (n=750)) Its nonlinear spline basis allows more flexibility to approximate the curved surface.
- **Trigonometric Function:**
  Best Model: **Beta-Spline Regression** *Best Performance:* RMSE = 0.0359, MAE = 0.0294 (at (n=750)) Splines provide the necessary adaptability to capture periodic structure, whereas other models systematically fail.

Table (1): Summary of Predictive Accuracy Across Simulation Scenarios

Function	S.S	Model	RMSE_mean	MAE_mean
Exp-Log	250	Beta-Spline	0.0202	0.0160
Exp-Log	250	Classical-Beta	0.0154	0.0122
Exp-Log	250	Quantile_0.5	0.0234	0.0189
Exp-Log	250	Trans-Gaussian	0.0221	0.0185
Exp-Log	500	Beta-Spline	0.0141	0.0112
Exp-Log	500	Classical-Beta	0.0128	0.0100
Exp-Log	500	Quantile_0.5	0.0183	0.0149
Exp-Log	500	Trans-Gaussian	0.0188	0.0162
Ex-Log	750	Beta-Spline	0.0113	0.0090
Exp-Log	750	Classical-Beta	0.0117	0.0091
Exp-Log	750	Quantile_0.5	0.0162	0.0132
Exp-Log	750	Trans-Gaussian	0.0174	0.0152
Polynomial	250	Beta-Spline	0.0762	0.0582
Polynomial	250	Classical-Beta	0.0830	0.0653
Polynomial	250	Quantile_0.5	0.0768	0.0572
Polynomial	250	Trans-Gaussian	0.0787	0.0590
Polynomial	500	Beta-Spline	0.0733	0.0561
Polynomial	500	Classical-Beta	0.0821	0.0648
Polynomial	500	Quantile_0.5	0.0735	0.0547
Polynomial	500	Trans-Gaussian	0.0765	0.0573
Polynomial	750	Beta-Spline	0.0723	0.0552
Polynomial	750	Classical-Beta	0.0818	0.0645
Polynomial	750	Quantile_0.5	0.0727	0.0540
Polynomial	750	Trans-Gaussian	0.0761	0.0568
Trigonometric	250	Beta-Spline	0.0415	0.0329
Trigonometric	250	Classical-Beta	0.1785	0.1519
Trigonometric	250	Quantile_0.5	0.1109	0.0914
Trigonometric	250	Trans-Gaussian	0.1070	0.0870
Trigonometric	500	Beta-Spline	0.0370	0.0299
Trigonometric	500	Classical-Beta	0.1776	0.1514
Trigonometric	500	Quantile_0.5	0.1076	0.0895
Trigonometric	500	Trans-Gaussian	0.1047	0.0858
Trigonometric	750	Beta-Spline	0.0359	0.0294
Trigonometric	750	Classical-Beta	0.1776	0.1515
Trigonometric	750	Quantile_0.5	0.1069	0.0891
Trigonometric	750	Trans-Gaussian	0.1043	0.0855

Overall Trends:

1. **Spline-based Beta Regression is the most robust across all nonlinear settings**, performing best in the Polynomial and Trigonometric scenarios and performing nearly as well as the top method in the Exponential–Log case.
2. **Classical Beta Regression excels only when the true mean is close to a linear predictor transformation**, as in the Exponential–Log mechanism. It deteriorates sharply when presented with strong nonlinearity.
3. Sample size benefits all models but does not change relative rankings, indicating that model flexibility, rather than sample size, is the primary determinant of performance.

Visual Comparison of Model Performance:

To complement the numerical results presented in Table 1., Figures 1–3 illustrate the distribution of Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) obtained from 500 Monte Carlo replications for each combination of test function, sample size, and modeling approach. The boxplots provide a visual summary of predictive accuracy, variability, and robustness across the four competing methods.

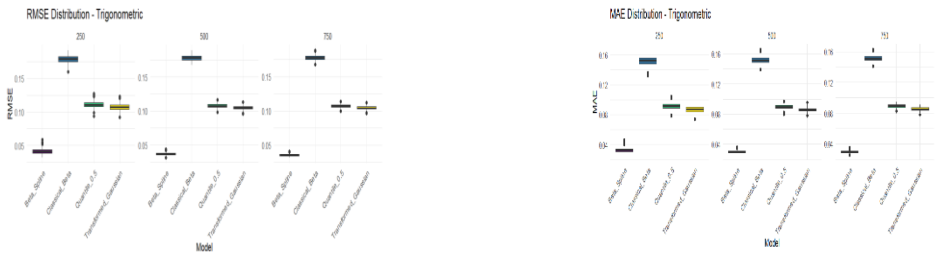

Figure (1): Distribution of RMSE (left) and MAE (right) for the Trigonometric test function across sample sizes n=250,500,750. Models: Spline Beta, Classical Beta, Quantile Regression, Transformed Gaussian.

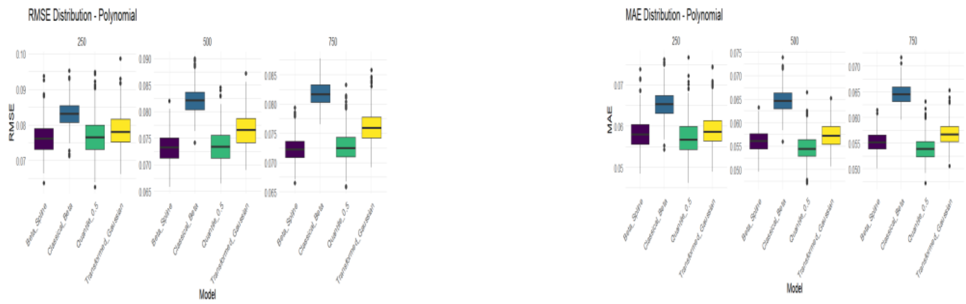

Figure (2): Distribution of RMSE (left) and MAE (right) for the Trigonometric test function across sample sizes n=250,500,750. Models: Spline Beta, Classical Beta, Quantile Regression, Transformed Gaussian

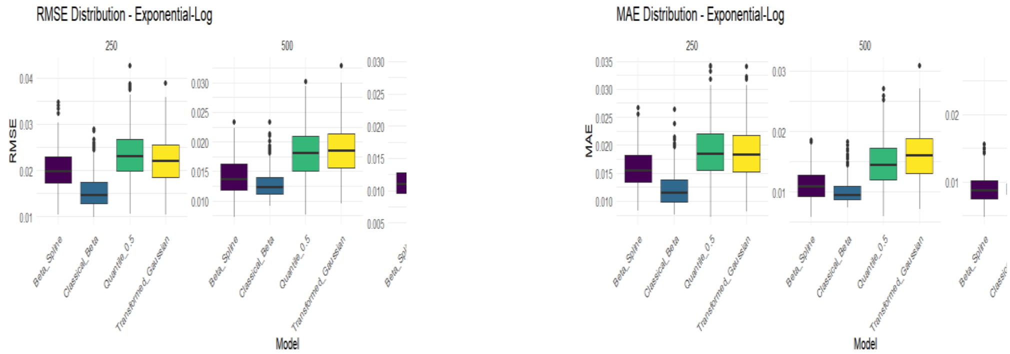

Figure (3)*:* Distribution of RMSE (left) and MAE (right) for the Trigonometric test function across sample sizes n=250,500,750. Models: Spline Beta, Classical Beta, Quantile Regression, Transformed Gaussian

Overall Patterns:

Spline-based beta regression exhibits the narrowest interquartile ranges (IQRs) and lowest median errors in nearly all settings, especially for the Trigonometric and Polynomial data-generating mechanisms. This indicates that the method not only achieves high point-wise accuracy but also delivers stable predictions across repeated samples.

Classical beta regression performs well under the Exponential-Log scenario where the mean structure is approximately linear on the logit scale but shows markedly wider and higher boxplots for the Trigonometric and Polynomial functions. Its inability to adapt to nonlinear shapes is visually evident in the elevated and dispersed error distributions.

Quantile regression and transformed Gaussian regression occupy an intermediate position. Their boxplots are generally broader than those of the spline-based approach, reflecting greater simulation-to-simulation variability. In the Trigonometric setting, both methods show substantial overlap in error distributions, yet neither approaches the low median error of the spline-based model.

Effect of Sample Size:

Increasing the sample size from n=250, n=500, and n=750 systematically reduces the spread of the boxplots for all models, as expected. However, the relative ordering of the methods remains unchanged: the spline-based beta regression continues to outperform the alternatives at every sample size. This reinforces the conclusion that model flexibility, rather than sheer data quantity, is the primary determinant of performance when the underlying mean and precision surfaces are nonlinear.

Robustness to Different Functional Forms:

The figures highlight how each method responds to qualitatively different types of nonlinearity: Trigonometric function: Only the spline-based beta regression maintains a tight, low-lying error distribution; the other models produce widely dispersed, high-median boxplots.

Polynomial function: The spline-based and quantile-regression boxplots are closely aligned and distinctly lower than those of the classical beta and transformed Gaussian models.

Exponential-Log function: All models yield compact boxplots, with classical beta regression slightly edging out the spline-based version in median error a result consistent with the near-linear nature of the generating process.

Discussion:

Our simulation study reveals clear trade-offs among competing methods for modeling bounded continuous outcomes. The proposed spline-based beta regression consistently outperformed traditional approaches when the data exhibited strong nonlinearity, effectively capturing complex patterns in both the mean and precision without sacrificing interpretability. This balance between flexibility and clarity is a key strength splines allow the model to “bend” where needed while still providing a structured, understandable form. In contrast, the transformed Gaussian model while intuitive and widely used struggles with two fundamental issues. First, it models the data on a transformed scale, making back-transformation to the original mean nontrivial and often biased. Second, it cannot accommodate varying precision across observations, a common reality in bounded data. These limitations make it less reliable in practice, despite its computational simplicity. Quantile regression offers robustness and avoids distributional assumptions, making it appealing for median estimation and outlier-resistant inference. Yet, it falls short when the goal is full probabilistic modeling or mean prediction, since estimating multiple quantiles separately can be inefficient and may even lead to logically inconsistent results (e.g., quantile crossing). For researchers interested in the full conditional distribution not just specific quantiles a likelihood-based approach remains preferable. As expected, larger sample sizes helped all models, reducing variability and sharpening estimates. Even with more data, a poorly specified model (like the classical linear beta regression) still struggled with trigonometric patterns, while the spline model adapted gracefully.

Conclusion

The proposed framework thus bridges an important gap in the methodological toolkit for bounded response modeling. It offers a flexible alternative to traditional linear models and provides greater interpretability compared to other approaches. Future research could explore the incorporation of regularization techniques, interaction effects, or spatially and temporally correlated structures within this spline-based framework. In summary, the spline-enhanced beta regression model represents a robust, adaptable, and interpretable choice for researchers and practitioners working with bounded data, supporting both accurate inference and clear communication of results in applied settings.

Reference:

1. Ferrari, S. L. P., & Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. Journal of Applied Statistics, 31(7), 799–815.
2. Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. Psychological Methods, 11(1), 54–71.
3. Ma,Y.,Li,S.&Zhang,Y.(2019). Bayesian beta regression with covariant-dependent precision. Computational Statistics & Data Analysis, 135, 1-18.
4. Wood, S. N. (2017). Generalized additive models: An introduction with R (2nd ed.). CRC Press.
5. Wang, X., Li, Y., & Zhang, Z. (2020). Neural network approaches for modeling bounded continuous outcomes. [Journal/Conference information needed].
6. Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. Journal of Applied Econometrics, 11(6), 619–632.
7. Taylor, J. W. (2019). Quantile regression for bounded data. Journal of Applied Statistics, 46(8), 1401–1416.
8. Figueroa-Zúñiga, J. I., Arellano-Valle, R. B., & Ferrari, S. L. P. (2013). Mixed beta regression: A Bayesian perspective. Computational Statistics & Data Analysis, 61, 137–147.
9. Meaney, C., & Moineddin, R. (2014). A Monte Carlo simulation study comparing linear regression, beta regression, variable-dispersion beta regression and fractional logit regression at recovering average difference measures in a two-sample design. BMC Medical Research Methodology, 14, 14.