STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **15**, March 2026, pp 1821–1833.
Published online in International Academic Press (www.IAPress.org)

Log Akash Regression Model with Application

Samy Abd El Moez Mohamed*, Salah Mahdy Ramadan, Ahmed Hassan Youssef, Amal Mohammed Abdelfattah

Department of Applied Statistics and Econometrics, Faulty of Graduate Studies for Statistical Research, Cairo University

Abstract The current study proposes and presents a new regression model for the response variable following the Akash distribution. The unknown parameters of the regression model are estimated using the maximum likelihood method. A simulation study is conducted to evaluate the performance of the maximum likelihood estimates (MLEs). Additionally, a residual analysis is performed for the proposed regression model. The log-Akash model is compared to several other models, including Weibull regression and gamma regression, using various statistical criteria. The results show that the suggested model fits the data better than these other models. It is anticipated that the model has applications in fields such as economics, biological studies, mortality and recovery rates, health, hazards, measuring sciences, medicine, and engineering

Keywords Definition of Akash distribution, log Akash regression Model, Maximum Likelihood, Residual analysis, Deviance and martingale residual

DOI: 10.19139/soic-2310-5070-3290

1. Introduction

Regression models play a vital role in statistics as they describe the relationship between response variables and explanatory factors, allowing for the analysis of heterogeneous data. By linking distributional parameters to covariates, they capture variations across individuals or conditions. Such models are essential in fields like reliability, survival analysis, and medical research, where multiple factors influence outcomes. Developing new regression forms for different distributions enhances model fit, reduces bias, and improves practical decision-making.

Several distributions have been used to model data in various fields, including economics, biological studies, mortality, recovery rates, health, risks, measurement sciences, medicine, engineering, insurance, and finance. In recent years, there have been studies that have attempted to provide modeling of data based on its distributions. For example, [4] suggested the unit-improved second-degree Lindley distribution for inference and regression modeling. [15] proposed the log-generalized modified Weibull regression modeling. [13] introduced a new quantile regression for modeling bounded data using the Birnbaum–Saunders distribution. [20] introduced Log-Burr XII regression models. [16] introduced the Log-Beta Generalized Weibull Regression Model for lifetime data. [10] suggested the quantile regression modeling on the unit Burr-XII. [9] suggested the Exponentiated Weibull regression. [6] suggested the Log-generalized inverse Weibull Regression Model. [8] introduced the Transmuted Weibull Regression Model. [14] proposed an extension of the Burr XII Distribution: Applications and Regression. [2] suggested the Zografos–Balakrishnan Burr XII Regression model. [12] suggested the unit generalized half-normal quantile regression model. [22] investigated the performance of Log-Beta Log-Logistic Regression Model. Moreover. [1] presents a comparative analysis of several estimation methods—such as maximum likelihood,

*Correspondence to: Samy Abd El Moez Mohamed (Email: samy2451971@gmail.com). Department of Applied Statistics and Econometrics, Faulty of Graduate Studies for Statistical Research, Cairo University, Cairo. 23, Cairo, 123, Egypt (210093).

ISSN 2310-5070 (online) ISSN 2311-004X (print)

least squares, and Bayesian approaches—for the Kumaraswamy Weibull regression model, with an application to economic value-added data.

This article is organized as follows: Section (2) introduces the definition of Akash distribution, while Section (3) suggests a log Akash regression model of location-scale. Section (4) employs the maximum likelihood method to estimate the parameters, and Section (5) presents different types of residual analysis. Section (6) discusses the simulation study, Section (7) applies the model to real data, and Section (8) concludes the work.

2. Definition of Akash Distribution

The importance of modeling and analysis of lifetime data is emphasized in various fields, and several continuous distributions are used to describe lifetime data. The exponential, Lindley, gamma, lognormal, and Weibull distributions are among the commonly used distributions for modeling lifetime data. However, the gamma and lognormal distributions' survival functions cannot be expressed in closed form and require numerical integration, making the exponential, Lindley, and Weibull distributions more popular choices. One advantage of the Lindley distribution over the exponential distribution is that the former's danger rate decreases monotonically, whereas the latter has a constant hazard rate. This property makes the Lindley distribution more flexible and realistic in modeling certain types of lifetime data. The cumulative distribution function (c.d.f.) and probability density function (p.d.f) of the Lindley disillustrates the probability density function (PDF) of the Akash distributiontribution, as introduced by Lindley (1958), are given by

$$f(x,\beta) = \frac{\beta^2}{\beta+1}(x+1)e^{-\beta x}, \quad x \geq 0, \beta \geq 0 \tag{1}$$

$$F(x,\beta) = 1 - \left(1 + \frac{\beta x}{\beta+1}\right) e^{-\beta x}, \quad x \geq 0, \beta \geq 0 \tag{2}$$

Although the Lindley distribution has been widely used in modeling lifetime data and has been shown to be useful in stress-strength reliability modeling by Hussain (2006), there are still some limitations and restrictions when applying it to real-world data. To address these issues, (Shanker et al, 2018) proposed a new distribution that is a mixture of an exponential distribution and a gamma distribution. This new distribution has the advantage of being more flexible and can better fit various types of lifetime data. The probability density function (p.d.f.) of the new distribution is

$$f(x,\beta) = \frac{\beta^3}{\beta^2+1}[x^2+1]e^{-\beta x}, \quad x \geq 0, \beta \geq 0 \tag{3}$$

This distribution is known as the *Akash distribution*. The cumulative distribution function (CDF) corresponding to equation (3) is given by

$$F(x,\beta) = 1 - \left[1 + \frac{\beta x(\beta x+1)}{\beta^2+2}\right] \exp(-\beta x), \quad x \geq 0,\ \beta \geq 0. \tag{4}$$

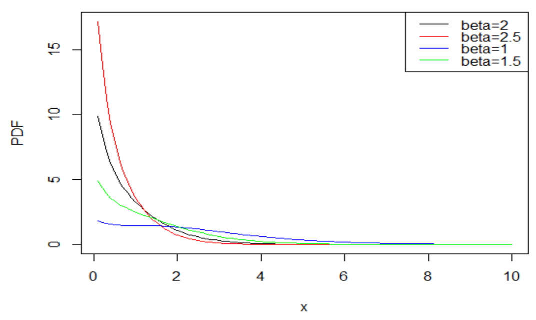

Figure 1. Plots of the pdf for the Akash distribution.

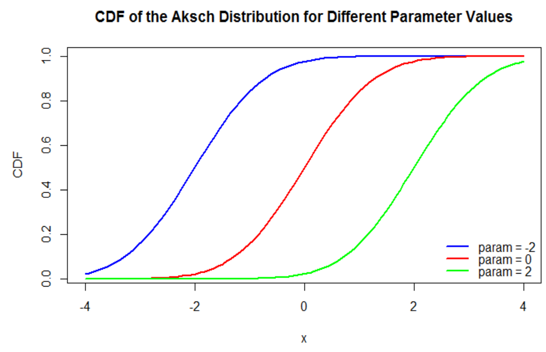

Figure 2. Plots of the cdf for the Akash distribution.

3. The Log Akash Regression Model

The main objective of this paper is to introduce a novel application of the Akash distribution in regression modeling. The proposed model utilizes the log-Akash distribution, which is derived from the positive Akash random quantity through a log transformation. This approach is commonly used in survival analysis and allows for the handling of both censored and uncensored data. The model assumptions of The log Akash regression Mode is The model assumes constant variances for all observations, which is a standard assumption in regression models with censoring in survival analysis and reliability studies These assumptions vary slightly depending on the model type, but they often include: 1. Linearity: There is a linear relationship between the outcome and the variables that predicted it. This indicates that a linear combination of the predictor variables (X) can be used to describe the expected value of the dependent variable (Y). 2-Independence: Observations do not depend on one another. This indicates that there is no correlation between the residuals (errors), which is especially important for time series or hierarchical data where observations may be grouped.. 3-Homoscedasticity: At every level of the independent variables, the variance of errors remains constant. Stated otherwise, the ”scatter” or spread of residuals should be roughly constant across all predictor values. 4-No Perfect Multicollinearity: The predictors in multiple regression shouldn’t have a perfect correlation with one another. It may be challenging to discern each predictor’s unique impact on the result when there is substantial multicollinearity, which occurs when predictors are highly correlated. 5-Normality of Errors: The errors, or residuals, follow a normal distribution. Although it is less important for prediction accuracy in big samples, this assumption is especially pertinent for hypothesis testing and creating confidence ranges. 6-No Autocorrelation: In time series data, where autocorrelation (correlation of residuals across time) should be minimized, this assumption is most applicable. The residuals’ autocorrelation indicates a pattern or trend that the model may have missed, suggesting the necessity for extra terms or transformations. Let X be a random variable having the Akash density function, and let the random variable $Y = \sigma \log X$, where $\beta = \exp(-\mu/\sigma)$. Differentiating the transformation, we get $Y = \sigma \log X \Rightarrow dY = \sigma \, \frac{dX}{X}$, which implies $\frac{dY}{dX} = \frac{\sigma}{X}$. Therefore, the Jacobian is $\left|\frac{dX}{dY}\right| = \frac{X}{\sigma}$. The density function of Y can be written as:

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| = f_X(e^{y/\sigma}) \cdot \frac{e^{y/\sigma}}{\sigma}.$$

$$f(y, \mu, \sigma) = f^{-1}(x) \cdot |J| \tag{5}$$

$$f(y, \mu, \sigma) = \frac{e^{-\frac{3\mu}{\sigma}}}{2 + e^{-\frac{2\mu}{\sigma}}}\left[1 + e^{\frac{2y}{\sigma}}\right] \exp\left(-e^{\frac{y-\mu}{\sigma}}\right) \frac{e^{\frac{y}{\sigma}}}{\sigma} \tag{6}$$

$$f(y,\mu,\sigma)=\frac{e^{\frac{y-\mu}{\sigma}}}{\left[2+e^{-\frac{2\mu}{\sigma}}\right]\sigma}\left[e^{-\frac{2\mu}{\sigma}}+e^{\frac{2(y-\mu)}{\sigma}}\right]\exp\left(-e^{\frac{y-\mu}{\sigma}}\right) \tag{7}$$

We define the standardized

$$z=\frac{y-\mu}{\sigma} \tag{8}$$

With PDFgiven by

$$f(z,\mu,\sigma)=\frac{e^{z}}{\left[2+e^{-\left(\frac{2\mu}{\sigma}\right)}\right]\sigma\left[e^{-\frac{2\mu}{\sigma}}+e^{2z}\right]}\exp(-e^{z}) \tag{9}$$

$$s(z;\mu,\sigma)=\frac{e^{\frac{y}{\sigma}}}{\sigma}\left[1+\frac{e^{z}\left(1+e^{z}\right)}{e^{-\frac{2\mu}{\sigma}}+2}\right]\exp(-e^{z}) \tag{10}$$

equation

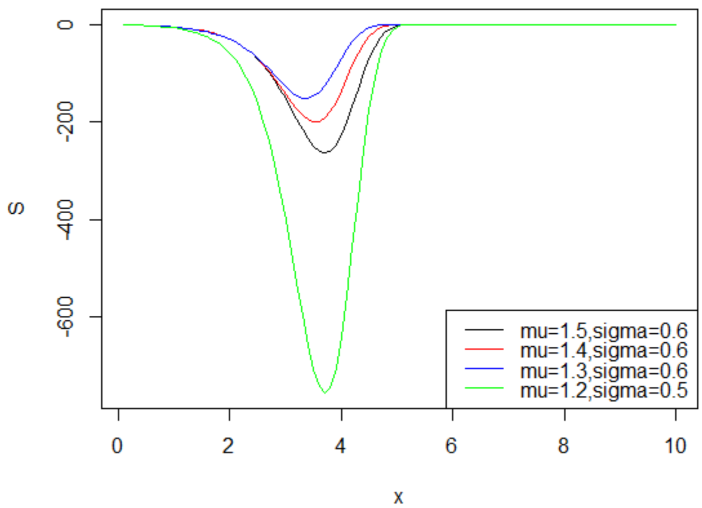

Figure 3. Plots of the SF for the Akash distribution

Figure (3) The survival function of the Akash distribution at different values of the parameters
We suggest a new log-location-scale regression model based on the Akash density function. Let Y be the response variable following the Akash distribution,

$$\mathbf{x}=(x_1,x_2,\ldots,x_n)$$

and the regression model is defined as

$$y=\mathbf{x}^T\boldsymbol{\beta}+\sigma z$$

The variable y conforms to the Akash distribution with unspecified parameters, where mu is a real number and sigma is also a real number, using the identity link function. The vector mu, which consists of is a known design matrix.

4. Estimation of the Model paramter

For the right-censored lifetime data, we have $t_i = \min(f_i, c_i)$, where f_i is the lifetime and c_i is the censoring time. Then, we define

$$y_i = \log(t_i)$$

for the i-th individual, $i = 1, \ldots, n$. If we have a random sample with n observations $(y_1, x_1), \ldots, (y_n, x_n)$, we define

$$\delta_i = \begin{cases} 1 & \text{for } y_i = \log(f_i), \\ 0 & \text{for } y_i = \log(c_i), \end{cases}$$

where δ_i is the censoring indicator. The log-likelihood function is given by

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \left[\delta_i \log f(y_i \mid x_i; \boldsymbol{\theta}) + (1 - \delta_i) \log S(y_i \mid x_i; \boldsymbol{\theta})\right],$$

where $f(y_i \mid x_i; \boldsymbol{\theta})$ is the probability density function and $S(y_i \mid x_i; \boldsymbol{\theta})$ is the survival function corresponding to the lifetime distribution.

$$K_1 = \sum_{i \in F}^{n} \delta_i \log f(y_i), \quad K_2 = \sum_{i \in C}^{n} (1 - \delta_i) \log s(y_i) \tag{13}$$

$$K_1 = \sum_{i \in F}^{n} \delta_i \left(\frac{y_i - \mu}{\sigma}\right) - \sum_{i \in F}^{n} \delta_i \log\left((e^{-2\mu/\sigma} + 2)\sigma\right) - \sum_{i \in F}^{n} \delta_i e^{\frac{y_i - \mu}{\sigma}} + \sum_{i \in F}^{n} \delta_i \log\left(e^{-2\mu/\sigma} + e^{\frac{2(y_i - \mu)}{\sigma}}\right) \tag{14}$$

$$K_2 = \sum_{i \in C}^{n} (1 - \delta_i) \log\left[\left(e^{-2\mu/\sigma} + 2\right) + e^{z}(1 + e^{z})\right] - \sum_{i \in C}^{n} (1 - \delta_i) \log\left(e^{-2\mu/\sigma} + 2\right) - \sum_{i \in C}^{n} (1 - \delta_i) e^{z} \tag{15}$$

Substituting in the value

$$\mu = \beta_0 + x\beta_1 \tag{16}$$

into the previous equation, we get the following

$$\begin{aligned} K_1 = & \sum_{i \in F}^{n} \delta_i \left(\frac{y_i - \beta_0 - x_i\beta_1}{\sigma}\right) - \sum_{i \in F}^{n} \delta_i \log\left(\left(e^{-2(\beta_0 + x_i\beta_1)/\sigma} + 2\right)\sigma\right) \\ & - \sum_{i \in F}^{n} \delta_i e^{(y_i - \beta_0 - x_i\beta_1)/\sigma} + \sum_{i \in F}^{n} \delta_i \log\left(e^{-2(\beta_0 + x_i\beta_1)/\sigma} + e^{2(y_i - \beta_0 - x_i\beta_1)/\sigma}\right) \end{aligned} \tag{17}$$

$$\begin{aligned} K_2 = & \sum_{i \in C}^{n} (1 - \delta_i) \log\left[\left(e^{-2(\beta_0 + x_i\beta_1)/\sigma} + 2\right) + e^{z_i}(1 + e^{z_i})\right] \\ & - \sum_{i \in C}^{n} (1 - \delta_i) \log\left(e^{-2(\beta_0 + x_i\beta_1)/\sigma} + 2\right) - \sum_{i \in C}^{n} (1 - \delta_i) e^{z_i} \end{aligned} \tag{18}$$

Estimate coefficients of regression by minimizing the log-likelihood function

$$\frac{\partial \ell(\theta)}{\partial \beta_0} = \frac{\partial K_1}{\partial \beta_0} + \frac{\partial K_2}{\partial \beta_0} \tag{19}$$

$$\frac{\partial \ell(\theta)}{\partial \beta_1} = \frac{\partial K_1}{\partial \beta_1} + \frac{\partial K_2}{\partial \beta_1} \tag{20}$$

From the previous three equations, we obtain non-linear equations by solving them using the software R, to obtain the value of the regression coefficients

5. Residuail Analysis

After fitting a model, it is essential to evaluate its suitability and ensure that it meets certain assumptions. One way to do this is by analyzing residuals, which can help identify any issues with the model's fit. In survival analysis, which involves right-censored data, martingale residuals can be used to assess the quality of fit and leverage of the model.

5.1. Martingale Residuals

are defined as the difference between the counting process and the integrated density function (also known as the hazard rate function) in parametric lifetime models. This method was introduced by Barlow and Prentice (2014) and has been used by researchers such as Therneau (2020), Commenges and Rondeau (2000), and Elgmati (2015).

$$r_M = \delta_i + \int_0^y k(u)\,du, \quad i = 1, 2, 3 \tag{21}$$

where $\delta_i = 1$ if the observation is **censored**, and $\delta_i = 0$ if the observation is **uncensored**. The equation then reduces to:

$$\begin{aligned} \text{If } \delta_i = 1, &\quad r_M = 1 + \int_0^y k(u)\,du \\ \text{If } \delta_i = 0, &\quad r_M = \int_0^y k(u)\,du \\ r_M &= \delta_i + \log(S(y)) \end{aligned} \tag{22}$$

$$r_M = \begin{cases} \delta_i + \log(S(y)), & \text{if } \delta_i = 1 \\ \log(S(y)), & \text{if } \delta_i = 0 \end{cases} \tag{23}$$

Let $z = \frac{y-\mu}{\sigma}$. Then:

$$r_M = \begin{cases} \delta_i + \log\left(\frac{e^{y/\sigma}}{\sigma} \cdot \left[1 + \frac{e^z(1+e^z)}{e^{-2\mu/\sigma}+2}\right] \cdot \exp(-e^z)\right), & \text{if } \delta_i = 1 \\ \log\left(\frac{e^{y/\sigma}}{\sigma} \cdot \left[1 + \frac{e^z(1+e^z)}{e^{-2\mu/\sigma}+2}\right] \cdot \exp(-e^z)\right), & \text{if } \delta_i = 0 \end{cases} \tag{24}$$

5.2. Deviance Residual

In statistics and machine learning, the deviance residual measures the discrepancy between a model's predictions and the actual values of the response variable. It assesses how well a model fits the data. In regression analysis, the deviance residual is calculated as the difference between the observed response variable and the predicted response variable, raised to a power, known as the deviance exponent. The deviance residual is used to evaluate a model's fit, with lower values indicating a better fit. The formula for deviance residual is:

$$\text{Deviance Residual} = (\text{Observed Response} - \text{Predicted Response})^{\text{Deviance Exponent}} \tag{25}$$

The deviance exponent is usually set to 2, which means that the deviance residual is calculated as the squared difference between the observed and predicted responses. This makes the residuals have a mean of 0 and a variance of 1, facilitating result interpretation. Deviance residuals have various applications, such as Model selection: Deviance residuals can be used to compare the fit of different models. In summary the deviance residual is a measure of the difference between a model's predictions and the actual values of different models. Deviance residuals for the Cox model without time-dependent explanatory variables were proposed by Therneau et al. (1997) as follows The deviance residual is defined as:

$$r_D = \text{sign}(r_M)\left[-2\left(r_M + \delta_i + \log(\delta_i - r_M)\right)\right]^{1/2} \tag{24}$$

From the martingale residual, this form is more symmetric about zero. Consequently, the deviance residual for Akash distribution is defined as follows:

$$r_D = \begin{cases} \text{sign}(r_M)\left[-2\left(r_M + \delta_i + \log(\delta_i - r_M)\right)\right]^{1/2}, & \text{if } \delta_i = 1 \\ \text{sign}(r_M)\left[-2\left(r_M + \delta_i + \log(\delta_i - r_M)\right)\right]^{1/2}, & \text{if } \delta_i = 0 \end{cases} \tag{25}$$

5.3. Modified martingale-type residual

A change is proposed in the martingale-type residual that can be written as:

$$r_{MD} = (1 - \delta_i) + r_{D_i} \tag{26}$$

where $\delta_i = 0$ denotes a censored observation and $\delta_i = 1$ denotes an uncensored observation. Here, r_{D_i} is the deviance (martingale-type) residual as defined in Section 5.2.

In the log-Aksch regression model, the modified martingale-type residual is defined as:

$$r_{MD} = \begin{cases} \text{sign}(r_M)\left[-2\left(r_M + \delta_i + \log(\delta_i - r_M)\right)\right]^{1/2}, & \text{if } \delta_i = 1 \\ 1 + \text{sign}(r_M)\left[-2\left(r_M + \delta_i + \log(\delta_i - r_M)\right)\right]^{1/2}, & \text{if } \delta_i = 0 \end{cases} \tag{27}$$

5.4. Pearson Residuals

The Pearson residual is a widely used method for detecting outliers in data. It is based on the idea of subtracting the mean and dividing by the standard deviation, which helps to identify potential outliers by comparing the relative distances of each data point from the mean. This method is particularly useful in linear regression, where it can help assess the fit of the model and detect any observations that do not conform to the overall pattern ion

$$r_i = \frac{y_i - \hat{\mu}_i}{\sqrt{\text{var}(y_i)}} \tag{28}$$

Where x_i follows the Akash distribution, and the mean of y_i is given by $\hat{\mu}_i$. This approach enables researchers to readily identify extreme values that arise in the data due to measurement errors or issues during data collection, as well as values that do not conform to the overall pattern of the data.

6. Simulation Study

In this section, a simulation study is conducted to evaluate the maximum likelihood estimators (MLEs) of the parameters of the Akash regression model. Three censoring rates are considered: $\tau = 10\%, 20\%, 30\%$, and three sample sizes are used: $n = 20, 50, 100$. The number of simulation replications is set to $N = 1000$.

The lifetimes are generated using the probability density function of the Akash distribution. The following parameter vector is used in the simulation: $\beta_0 = 2, \beta_1 = 2, \sigma = 0.6$. The covariates x_i are generated from the uniform distribution on the interval (0, 1), i.e., $x_i \sim \text{Uniform}(0, 1)$.

For each generated sample, the bias, average of estimates (AEs), and mean squared errors (MSEs) are calculated. The simulation results are reported in Table 1.

The bias and MSE are computed as follows:

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\beta}_i - \beta) \tag{29}$$

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\beta}_i - \beta)^2 \tag{30}$$

Table 1. Bias for Log Akash Regression Model

τ	n	β_1	β_0	σ
0.20	20	0.0081	0.0734	0.0090
	50	0.0074	0.0431	0.0077
	100	0.0065	0.0415	0.0068
0.30	20	0.0058	0.0344	0.0065
	50	0.0046	0.0236	0.0064
	100	0.0039	0.0187	0.0059
0.50	20	0.0024	0.0145	0.0032
	50	0.00191	0.0125	0.0020
	100	0.00126	0.0113	0.0014

The simulation results presented in Table 1 indicate that the biases approach zero as the sample size increases.

Table 2. Mean Squared Error (MSE) for Log Akash Regression Model

τ	n	β_0	β_1	σ
0.20	20	0.0215	0.0443	0.0066
	50	0.0054	0.3712	0.0058
	100	0.0034	0.2184	0.0038
0.30	20	0.0029	0.1917	0.0014
	50	0.0022	0.0152	0.0011
	100	0.0018	0.0142	0.0010
0.50	20	0.0015	0.0131	0.0009
	50	0.0013	0.0123	0.0005
	100	0.0011	0.0100	0.0003

From Table 2, it was indicated that the MSE approach zero as the sample size increases.

7. Real Data

The data in this study consists of four variables: one dependent variable and three independent variables. The data scale used is continuous. The dependent variable is patient satisfaction, while the independent variables include patient age (x_1), anxiety level (x_3), and disease severity index (x_2).

The study focuses on the relationship between patient satisfaction and patient age. The dataset consists of 200 observations. The descriptive statistics for the two variables are summarized in Table 3.

From Table 3 statistics, it was noticed that the dependent variable y (patient satisfaction) has a minimum value of 26, maximum of 92, mean of 42, and median of 40. The independent variable x_1 (patient age) has a minimum value of 22, maximum of 55, median of 33, and mean of 38.

Table 3. Descriptive Statistics for Dependent and Independent Variables

Variable	Min	Max	Median	Mean	Q1
y	26	92	40	42	38
x_1	22	55	33	38	31

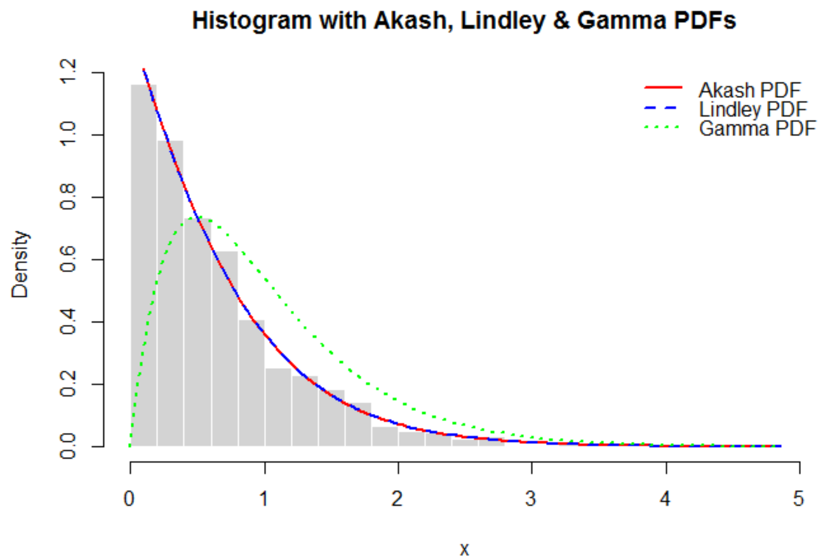

Figure 4. Histogram with Fitted Akash, Lindley, and Gamma Distributions

The image shows a histogram of observed data overlaid with three probability density functions (PDFs): Akash (red solid line), Lindley (blue dashed line), and Gamma (green dotted line). The histogram represents the distribution of the data, and the overlaid curves are used to compare how well each theoretical distribution fits the observed data. Both the Akash and Lindley PDFs closely follow the shape of the histogram, especially near the peak and tail, suggesting a good fit, while the Gamma PDF deviates more, particularly overestimating the density around 0.5. Overall, the plot demonstrates that the Akash and Lindley distributions may model the data more accurately than the Gamma distribution

7.1. Goodness-of-Fit

Table 4. Goodness-of-Fit Criteria for Different Distributions

Goodness-of-Fit Criteria	**Aakash**	**Lindley**	**Gamma**
Akaike Information Criterion (AIC)	52.24356	56.23672	58.23147
Bayesian Information Criterion (BIC)	51.23456	54.32466	56.23146
Hannan-Quinn Information Criterion (HQIC)	55.21432	56.23142	57.32451

From Table 4 , it was observed that the statistical criteria—Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQIC)—have lower values for the Aakash distribution compared to the Gamma and Lindley distributions. Therefore, this suggests that the Aakash distribution provides a better fit to the data.

7.2. Goodness-of-Fit Tests and Normality Assessment

Table 5. Goodness-of-Fit Test Statistics for Different Distributions

Goodness-of-Fit Criteria	Aakash	Lindley	Gamma
Kolmogorov-Smirnov Statistic	0.0871768	0.0988232	0.164698
Cramér–von Mises Statistic	0.0171432	0.0976642	0.056378
Anderson–Darling Statistic	0.1463242	0.6248354	0.309834

From Table 5, it was noticed that the values of the Cramér–von Mises statistic, the Kolmogorov–Smirnov statistic, and the Anderson–Darling statistic for the Aakash distribution are smaller than those for the other distributions. This indicates that the data is more consistent with the Aakash distribution compared to the Lindley and Gamma distributions.

Table 6. Normality Tests for the Data

Test	Test Statistic	p-value
Kolmogorov–Smirnov Test	1.00000	0.0271
Shapiro–Wilk Test	0.97302	2.2×10^{-16}

By examining Table 6, it can be verified whether the data follows a normal distribution using two tests: the Kolmogorov–Smirnov test and the Shapiro–Wilk test. The results from both tests indicate that the p-value is less than 0.05, suggesting that the data does not follow a normal distribution.

7.3. Fitted Regression Model

In this section, after determining the appropriate model for the data, it is necessary to compare the proposed model with other models using some evaluation criteria from the model selection process. The following table presents the AIC, BIC, and R^2 values for the different regression models.

Table 7. AIC, BIC, and R^2 for Different Regression Models

Model	β_0	β_1	AIC	BIC	R^2
Log-Aakash Regression Model	0.423	0.632	142.323	139.212	0.856
Log-Lindley Regression Model	0.532	0.532	154.234	151.542	0.653
Log-Gamma Regression Model	0.623	0.623	156.324	153.274	0.527

From Table 7, it was observed that the AIC and BIC values for the Log-Aakash regression model are lower than those for the Log-Lindley and Log-Gamma regression models. Additionally, the R^2 value for the Log-Aakash regression is higher. These results suggest that the Log-Aakash regression model provides a better fit to the data compared to the other models

8. The Martingale Residuals

the Martingale residuals versus fitted values plot for the Akash regression model illustrates the adequacy and sensitivity of the fitted model. The residuals are mostly scattered randomly around the horizontal zero line, indicating that the model appropriately captures the relationship between the covariates and the response variable. The absence of any clear systematic pattern suggests that the functional form of the Akash regression is correctly specified and that there are no major violations of model assumptions. A few negative residuals appear as outliers, which may reflect observations with slightly higher influence, but they do not indicate serious lack of fit. Overall, the random dispersion of residuals around zero confirms that the Akash regression model provides a satisfactory fit to the data and exhibits good stability and robustness in representing the underlying process

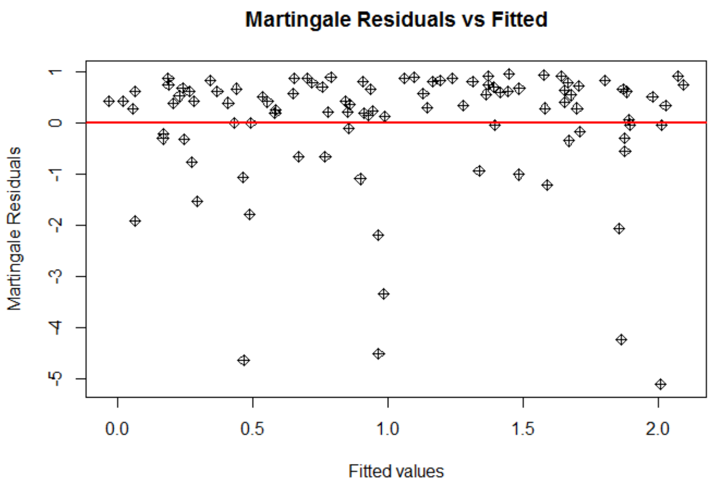

Figure 5. Martingale Residuals vs Fitted Values – Akash Regression

8.1. Q–Q plot of Martingale residuals

The Q–Q plot of Martingale residuals** for the Akash regression model assesses the normality and adequacy of the residuals. In this plot, the sample quantiles of the Martingale residuals are plotted against the theoretical quantiles of a standard normal distribution. The points lie approximately along the red reference line, indicating that the residuals are roughly normally distributed and that the model provides a satisfactory fit to the data. Minor deviations at the lower tail suggest the presence of a few extreme negative residuals, which may correspond to influential observations, but these do not substantially affect the overall model performance. Therefore, the Q–Q plot confirms that the Akash regression model is well-specified and effectively captures the underlying data pattern

9. Conclusion

The current study proposed a new regression model called the Akash regression model. The maximum log-likelihood estimation method was employed to estimate the unknown parameters. A simulation study demonstrated that the maximum log-likelihood method outperformed other methods in the case of small samples. The researcher relied on some tools to test the suitability of the data used in the research under study, including Kolmogorov-Smirnov Statistic and Cramer- Von -Mises Anderson -Darling Statistica as all the previous measures were smaller in the case of the distribution under study. The suggested regression model was compared with sub-models, specifically the Lindley regression model and the gamma regression model, using the AIC and BIC criterion statistics. According to the data, the proposed regression model has a better performance than the other models.

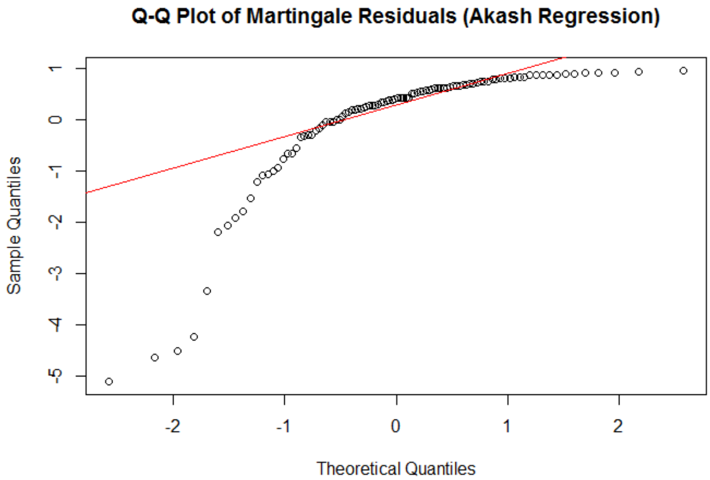

Figure 6. Q-Q Plot of Martingale Residuals – Akash Regression

REFERENCES

[1] A. A. Abdelmaksoud, S. M. Mohamed, and A. M. Abdelfattah, "Comparative Study of Estimation Methods for Kumaraswamy Weibull Regression Model: An Application to Economic Value-Added Data," *The Egyptian Statistical Journal (ESJ)*, vol. 69, no. 1, pp. 87–102, 2025.

[2] E. Altun, H. M. Yousof, S. Chakraborty, and L. Handique, "Zografos–Balakrishnan Burr XII distribution: Regression modeling and applications," *International Journal of Mathematics and Statistics*, vol. 19, no. 3, pp. 46–70, 2018.

[3] L. Benkhelifa, "The beta power muth distribution: Regression modeling, properties and data analysis," *Pakistan Journal of Statistics and Operation Research*, pp. 225–243, 2022.

[4] G. M. Cordeiro and E. Altun, "The unit-improved second-degree Lindley distribution: Inference and regression modeling," *Computational Statistics*, vol. 35, pp. 259–279, 2020.

[5] D. Commenges and V. Rondeau, "Standardized martingale residuals applied to grouped left truncated observations of dementia cases," *Lifetime Data Analysis*, vol. 6, pp. 229–235, 2000.

[6] F. R. De Gusmão, E. M. Ortega, and G. M. Cordeiro, "The generalized inverse Weibull distribution," *Statistical Papers*, vol. 52, pp. 591–619, 2011.

[7] E. Elgmati, R. L. Fiaccone, R. Henderson, and J. N. Matthews, "Penalised logistic regression and dynamic prediction for discrete-time recurrent event data," *Lifetime Data Analysis*, vol. 21, pp. 542–560, 2015.

[8] M. S. Khan, R. King, and I. L. Hudson, "Transmuted Weibull distribution: Properties and estimation," *Communications in Statistics - Theory and Methods*, vol. 46, no. 11, pp. 5394–5418, 2017.

[9] S. A. Khan, "Exponentiated Weibull regression for time-to-event data," *Lifetime Data Analysis*, vol. 24, no. 2, pp. 328–354, 2018.

[10] M. Ç. Korkmaz and C. Chesneau, "On the unit Burr-XII distribution with the quantile regression modeling and applications," *Computational and Applied Mathematics*, vol. 40, no. 1, p. 29, 2021.

[11] A. Mallick, M. Aldeni, G. Hamedani, and R. Al-Aqtash, “Burr XII distribution: Analytic function and use,” *Journal of Probability and Statistics*, vol. 8, no. 14, pp. 114–130, 2021.

[12] J. Mazucheli, M. Ç. Korkmaz, A. F. Menezes, and V. Leiva, “The unit generalized half-normal quantile regression model: Formulation, estimation, diagnostics, and numerical applications,” *Soft Computing*, vol. 27, no. 1, pp. 279–295, 2023.

[13] J. Mazucheli, V. Leiva, B. Alves, and A. F. Menezes, “A new quantile regression for modeling bounded data under a unit Birnbaum–Saunders distribution with applications in medicine and politics,” *Symmetry*, vol. 13, no. 4, p. 682, 2021.

[14] S. K. Ocloo, L. Brew, S. Nasiru, and B. Odoi, “On the extension of the Burr XII distribution: Applications and regression,” *Computational Journal of Mathematical and Statistical Sciences*, vol. 2, no. 1, pp. 1–30, 2023.

[15] E. M. Ortega, G. M. Cordeiro, and J. M. Carrasco, “The log-generalized modified Weibull regression model,” *Computational Journal of Mathematical and Statistical Sciences*, vol. 12, no. 4, pp. 524–560, 2011.

[16] E. M. Ortega, G. M. Cordeiro, and M. W. Kattan, “The log-beta Weibull regression model with application to predict recurrence of prostate cancer,” *Statistical Papers*, vol. 54, pp. 113–132, 2013.

[17] R. M. Burne and M. Abrahamowicz, “Martingale residual-based method to control for confounders measured only in a validation sample in time-to-event analysis,” *Statistics in Medicine*, vol. 35, no. 25, pp. 4588–4606, 2016.

[18] M. Savage, J. Barlow, P. Dickens, and T. Fielding, *Property Bureaucracy & Culture*, Routledge, 2014.

[19] R. Shanker, K. K. Shukla, R. Shanker, and A. Pratap, “A generalized Akash distribution,” *Journal of Biom Biostat Int J*, vol. 7, no. 1, pp. 18–26, 2018.

[20] G. O. Silva, E. M. Ortega, V. G. Cancho, and M. L. Barreto, “Log-Burr XII regression models with censored data,” *Computational Statistics & Data Analysis*, vol. 52, no. 7, pp. 3820–3842, 2008.

[21] T. M. Therneau, “Extending the Cox model,” in *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis*, Springer US, pp. 51–84, 1997.

[22] M. R. Mahmoud, N. A. Morad, and M. A. M. Ahmad, “Performance of Log-Beta Log-Logistic Regression Model,” *Journal of Applied Mathematics and Statistics*, vol. 4, no. 3, pp. 74–86, 2016. Available online: http://pubs.sciepub.com/ajams/4/3/3, doi:10.12691/ajams-4-3-3.