

A FLEXIBLE COUNT REGRESSION MODEL BASED ON POISSON-JUCHEZ DISTRIBUTION

Koyejo, S. O., *Akomolafe, A. A., Oseni, B. M. and Ojo, O. O.



DOI10.51459/jostir.2026.2.2.0342

Department of Statistics,
Federal University of
Technology, Akure

Correspondence

koyejosamuel@yahoo.co.uk,
akomolafeayotade@gmail.
com

History

Received: 19-06-2026

Accepted: 21-07-2026

Published: August, 2026

ABSTRACT

A new count regression model known as Poisson-Juchez regression (PJR) model is proposed in this study. The model is based on the Poisson-Juchez Distribution and is aimed to serve as a suitable count regression model. Its theoretical framework, estimation technique and simulation results were being studied. The proposed Poisson-Juchez regression model was being compared to traditional Poisson regression and Negative Binomial regression models to demonstrate its flexibility and applicability to real life data. The goodness of fit test showed that Poisson-Juchez regression model was better than both the standard Poisson and Negative Binomial models as demonstrated by the model selection criteria using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) as model selection Metrics. The result for the dataset showed that Poisson-Juchez regression model achieved the best AIC (200.81) and Bayesian Information Criterion BIC (210.54), thus outperforming both the standard Poisson (AIC = 203.26, BIC = 212.98) and Negative Binomial (AIC = 202.79, BIC = 215.75). The result from the real-life dataset showed that the Poisson-Juchez regression model is the preferred model, as it had the least AIC and BIC. The bias and Mean Square Error (MSE) declined consistently with increasing sample sizes, indicating the estimator's consistency. The bias of β_0 was -0.1395 , whereas MSE was 0.3273 , when $n = 10$, while the maximum MSE value was 1.0344 for β_2 as negative coefficients add extra variability in smaller sample sizes. All biases were minimal (less than 0.004), and the MSE was minimized significantly (approximately 0.001) for all parameters at $n = 1000$. The finding from this study had substantiated that pre-pregnancy weight (lwt) was a significant predictor of prior preterm labour, and this was being upheld across all three models, thus establishing it as the key determinant of prior preterm labour frequency in this study.

Keywords: Count Regression, Poisson-Juchez, Overdispersion, Index of Dispersion, Pre-term labour



<https://www.futa.edu.ng>



<https://jostir.futa.edu.ng>

1. | Introduction

The traditional Poisson regression is characterized by the equality of mean and variance of the count outcome known as equidispersion. This property most often, is void in practice. The violation of this basic assumption is attributed to a situation where variance is greater

than mean, a phenomenon called overdispersion (Cameron & Trivedi, 2013). Overdispersion can be caused by unobserved heterogeneity and introducing random effects can address it (Breslow & Clayton, 1993). Tests of hypotheses in over dispersed Poisson regression and other quasi-likelihood models were studied by Breslow (1990). Regression-based tests for

overdispersion in the Poisson model also carried out by Cameron & Trivedi (1990). The Poisson framework is being extended by incorporating covariates through the Generalized Linear Mixed Models (GLMMs) platform (McCulloch *et al.*, 2011). Its applications cover a wide range of studies and areas, including hospital utilization modelling (Deb & Trivedi, 1997), accident counts analysis (Lord & Mannering, 2010), ecological studies (Bolker *et al.*, 2009), insurance claims frequency prediction (Denuit *et al.*, 2007), etc.

Poisson regression models are often applied to Count data analysis, where the behavior of the response (dependent) variable Y may be explained by predictor (independent or explanatory) variables X_{ij} , that is. Expectation of the response variable Y , conditional to the predictor variables X_{ij} . Such models are said to have justified the mean and variance equality assumption of the response variable, conditional to the predictor variables. Nevertheless, Poisson regression models are inadequate to address the phenomenon of overdispersion which is feasible in numerous real-world count data (Cameron and Trivedi, 2013, Zeviani *et al.*, 2014, Fávero and Belfiore, 2019). This is a scenario in which the variance of the response variable, conditional to the predictor variables is greater than the corresponding mean (Payne *et al.*, 2017, Dupuy, 2018). Overdispersion is generally caused by various factors, including heterogeneity in sample values, presence of outliers, correlated variables, omission of relevant predictors or zero inflation (Payne *et al.*, 2018). The Count data regression models are utilized for situations whose observations under study are non-negative integers. The relative performance of Poisson and negative binomial regression estimators was studied (Blackburn, 2020). Zeileis *et al.*, (2008) provided R Journal of Statistical software for count data regression Models. Over the years, new count regression models for mixed distributions have been developed, few of which are hereby outlined. Altun

(2018) proposed a new zero-inflated regression model with application in Actuarial Science. Altun (2019) further proposed a model for over-dispersed count data called Poisson quasi-Lindley regression model, Klakattawi *et al.*, (2018): a simple and adaptive dispersion regression model for count data, Gelman (2006): a data analysis procedure using Regression and Multilevel /Hierarchical models for social research, Altun (2020): a new one-parameter discrete distribution with associated regression and integer-valued autoregressive models. Rudhakumari *et al.*, (2022): a new count regression model based on the Poisson-Mirra distribution. Koyejo *et al.*, (2026), in a bid to address the overdispersion property of count data with reference to maternity lengths of stay, developed a Poisson-mixed distribution called Poisson-Juchez distribution as a more flexible alternative to some already existing Poisson-mixed distributions in handling overdispersion in count data modelling. Thus, this study proposes the development of the regression model associated with the novel Poisson-Juchez distribution called Poisson-Juchez regression model which is expected to be a more flexible model to address overdispersion in count data modelling which real-world count data exhibit.

2. | Materials and Methods

A Poisson model assumes that the error term follows a Poisson distribution and that the natural log of the response variable (Y), that is, $\ln(Y_i)$ is being modelled as a linear function of the coefficients, instead of modeling Y as a linear function of the regression coefficients. The general regression model for count data is expressed as:

$$\ln(Y_i) = \ln(\mu_i) = \beta_0 + \beta_1 X_{i1} + \dots + \beta_{ij}, \quad i = 1, 2, 3, \dots \quad (1)$$

where: μ_i represents the expected number of occurrences or incidence rate ratio of the phenomenon under study for a given fixed time-

interval,

β_0 is the intercept,

β_j is the coefficient estimated for each predictor variable X_j and i represents each observation in the sample. The Probability Mass Function (PMF) of the Poisson-Juchez mixed

$$f(x, \theta) = \frac{\theta^4}{(\theta^3 + \theta^2 + 6)} \left[\frac{(1+\theta)^3 + x(1+\theta)^2 + (1+\theta)^2 + x^3 + 6x^2 + 11x + 6}{(1+\theta)^{x+4}} \right] \quad (2)$$

Its Mean and Variance are derived via Moments as follows:

The r^{th} moment is defined as:

$$\mu_r = \frac{\theta^4}{(\theta^3 + \theta^2 + 6)(1+\theta)^4} \sum_{x=0}^{\infty} x^r \left[\frac{(1+\theta)^3 + x(1+\theta)^2 + (1+\theta)^2 + x^3 + 6x^2 + 11x + 6}{(1+\theta)^x} \right] \quad (3)$$

$$\text{Let } A = \frac{\theta^4}{(\theta^3 + \theta^2 + 6)(1+\theta)^4}$$

Let $r=1, 2$ to derive the first and second Moments, as given by equations (4) and (5), respectively.

First Moment $r=1$

$$\begin{aligned} \mu_1 &= A \left[(1+\theta)^3 \sum_{x=0}^{\infty} x(1+\theta)^{-x} + (1+\theta)^2 \sum_{x=0}^{\infty} x^2(1+\theta)^{-x} + (1+\theta)^2 \sum_{x=0}^{\infty} x(1+\theta)^{-x} \right. \\ &\quad \left. + \sum_{x=0}^{\infty} x^4(1+\theta)^{-x} + 6 \sum_{x=0}^{\infty} x^3(1+\theta)^{-x} + 11 \sum_{x=0}^{\infty} x^2(1+\theta)^{-x} + 6 \sum_{x=0}^{\infty} x(1+\theta)^{-x} \right] \\ \mu_1 &= \frac{1}{(\theta^3 + \theta^2 + 6)(\theta + 1)^3} \left[\theta^2(1+\theta)^3 + (1+\theta)^2(\theta + 2)\theta + \theta^2(1+\theta)^2 + \frac{(\theta + 2)(\theta^2 + 12\theta + 12)}{\theta} + 6(\theta^2 + 6\theta + 6) + 11(\theta + 2) + 6\theta^2 \right] \\ \mu_1 &= \frac{\theta^3 + 2\theta^2 + 24}{\theta(\theta^3 + \theta^2 + 6)} \end{aligned} \quad (4)$$

Second Moment $r = 2$

$$\begin{aligned} \mu_2 &= \frac{\theta^4}{(\theta^3 + \theta^2 + 6)(1+\theta)^4} \sum_{x=0}^{\infty} x^2 \left[\frac{(1+\theta)^3 + x(1+\theta)^2 + (1+\theta)^2 + x^3 + 6x^2 + 11x + 6}{(1+\theta)^x} \right] \\ \mu_2 &= \frac{(\theta^4 + 4\theta^3 + 6\theta^2 + 24\theta + 120)}{\theta^2(\theta^3 + \theta^2 + 6)} \end{aligned} \quad (5)$$

$$\text{Variance } \sigma^2 = \mu_2 - \mu_1^2 \quad (6)$$

$$\sigma^2 = \frac{(\theta^4 + 4\theta^3 + 6\theta^2 + 24\theta + 120)}{\theta^2(\theta^3 + \theta^2 + 6)} - \left(\frac{\theta^3 + 2\theta^2 + 24}{\theta(\theta^3 + \theta^2 + 6)} \right)^2$$

$$\sigma^2 = \frac{(\theta^4 + 4\theta^3 + 6\theta^2 + 24\theta + 120)}{\theta^2(\theta^3 + \theta^2 + 6)} - \frac{(\theta^3 + 2\theta^2 + 24)^2}{\theta^2(\theta^3 + \theta^2 + 6)^2}$$

$$\sigma^2 = \frac{(\theta^7 + 4\theta^6 + 6\theta^5) + 32\theta^4 + 120\theta^3 + 60\theta^2 + 144\theta + 144}{\theta^2(\theta^3 + \theta^2 + 6)^2} \quad (7)$$

The regression model defining the relationship between the response variable and the predictor variables is given by:

$$\log(\theta) = \beta_0 + \beta_{1x_1} + \beta_{2x_2} + \dots + \beta_{kx_k} \quad (8)$$

where:

Θ is the rate parameter of the Poisson distribution

$\beta_0, \beta_1, \dots, \beta_k$, are the regression coefficients

$X_1, X_2, \dots, X_k$, are the predictor variables

The likelihood function is given by:

The likelihood function is given by:

$$L(\beta, \theta) = \prod [\Sigma [P(y_i/\lambda_i) * J(y_i/\lambda_\theta)]] \quad (9)$$

where:

y_i is the observed count data

λ_i is the rate parameter of the Poisson distribution for the i th observation

θ is the substitution rate parameter of the Juchez distribution

We consider a regression structure for both the mean and dispersion parameter of the mixed Poisson-Juchez model. The Methodology for forming a Regression Model for Poisson-Juchez Distribution is given as follows:

Case 1: Developing the Regression model based on Mean

The derived probability density function (pdf) of the Poisson-Juchez distribution is given as:

$$f(x, \theta) = \frac{\theta^4[(1+\theta)^3 + x(1+\theta)^2 + (1+\theta)^2 + x^3 + 6x^2 + 11x + 6]}{(\theta^3 + \theta^2 + 6)(1+\theta)^{x+4}} \quad (10)$$

where:

$x = 0, 1, 2, \dots$ is the discrete random variable,

$\theta > 0$ is the parameter of the distribution.

The mean (μ_1) of the Poisson-Juchez distribution was derived as:

$$\mu_1 = \frac{\theta^3 + 2\theta^2 + 24}{\theta(\theta^3 + \theta^2 + 6)} \quad (11)$$

Rearranging to express Θ in terms of μ_1 , that is, rewrite Θ as a function of μ_1 involves solving the nonlinear equation

$$\mu_1 = \frac{\theta^3 + 2\theta^2 + 24}{\theta(\theta^3 + \theta^2 + 6)}$$

Since the Mean is in closed form, a Regression Model Linking μ_1 to Predictors is thus formed by modelling the mean μ_1 as a function of predictors X_i for observation i :

$$\mu_i = h(X_i, \beta) = \exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}) \quad (12)$$

The exponential link ensures positivity of μ_i .

$\beta_0, \beta_1, \dots, \beta_p$ are regression parameters to be estimated.

Case 2: Developing the Regression model based on the Parameter Θ

$$L(\theta) = \prod_{i=1}^n f(\theta, y) \quad (13)$$

$$L(\theta) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} \quad (14)$$

Equivalently,

$$\theta = \exp(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}) \quad (15)$$

$$\theta = \exp(X^T \beta)$$

3. | Results

Prior Pre-term Labour data set, obtained from MASS: birthwt (Venables & Ripley, 2002) was used for this study. The response variable is the number of previous preterm labours (ptl) with counts: 0–3, Predictors are maternal age in years (age) and maternal pre-pregnancy weight (lwt) in pounds. The data was originally collected through a study conducted at Baystate Medical Center in 1986, to model the frequency of prior preterm births as a function of continuous maternal demographic variables. Its descriptive Statistics is presented in Table 1.

Table 1 | Descriptive Statistics: Prior Preterm Labours (ptl ~ age + lwt)

Statistic	Value
Sample size (n)	189
Mean	0.1958
Variance	0.2434
Variance / Mean ratio	1.2432
Range	0 – 3
Zero counts	159 (84.1%)

Table 1 shows that the prior preterm labour (ptl) variable has a very low mean (0.196), with 84.1% of women reporting no prior preterm labours. The variance (0.243) slightly exceeds the mean,

producing a variance-to-mean ratio of 1.24. While this reflects only mild overdispersion, the extreme concentration of zeros may still undermine the standard Poisson assumption, thus motivating evaluation of the Negative Binomial and Poisson-Juchez models. R Programming Language was used for required and necessary computations.

3.1 | Simulation Results for Poisson-Juchez Regression Estimator

Monte Carlo simulation study was performed to determine the finite-sample properties of the Poisson Juchez Regression (PJR) estimator. The Monte Carlo simulation was run for sample sizes (n) ranging from 10 to 1000, with model dimension $p=2$ predictors. In addition, three different coefficient configurations were tested for each case. The evaluation of performance was done on five main criteria: (i) convergence ratio (Conv%), (ii) bias in estimating each coefficient, (iii) mean square error (MSE), (iv) coverage probability of 95% confidence interval (Cov%), and (v) computation time in seconds. In each simulation scenario, 200 replications were applied.

3.2 | Simulation Results for two- model dimension $p=2$ predictors

Tables 2., 3, and 4 present the simulation results for the two-model dimension ($p=2$ predictors) of the Poisson Juchez Regression model across the following three coefficient scenarios:

Scenario 1: Moderate Effects ($\beta_0 = 0.5$, $\beta_1 = 0.3$, $\beta_2 = -0.2$), Scenario 2: Strong Positive Effects ($\beta_0 = 1.0$, $\beta_1 = 0.8$, $\beta_2 = 0.5$), and Scenario 3: Mixed Large Effects ($\beta_0 = 0.2$, $\beta_1 = -0.6$, $\beta_2 = 0.9$)

Table 2 | Descriptive Statistics: Prior Preterm Labours (ptl ~ age + lwt)

n	Valid	Conv%	β_0 Bias	β_0 MSE	β_0 Cov%	β_1 Bias	β_1 MSE	β_1 Cov%	β_2 Bias	β_2 MSE	β_2 Cov%	Time(s)
10	200	100.0%	-0.1395	0.3273	98.0%	+0.0553	0.2776	97.0%	-0.1039	1.0344	96.5%	0.0472
20	200	100.0%	-0.1120	0.1042	95.0%	+0.0192	0.1066	91.5%	+0.0002	0.0919	92.5%	0.0733
30	200	100.0%	-0.0404	0.0520	97.0%	-0.0163	0.0428	95.5%	-0.0225	0.0473	95.5%	0.0961
50	200	100.0%	-0.0428	0.0327	94.0%	+0.0218	0.0277	96.5%	+0.0086	0.0292	94.5%	0.1533
100	200	100.0%	-0.0152	0.0144	93.5%	-0.0069	0.0129	95.5%	+0.0055	0.0140	93.0%	0.2832
250	200	100.0%	-0.0011	0.0046	96.5%	-0.0012	0.0052	93.5%	+0.0065	0.0048	95.5%	0.7316
500	200	100.0%	-0.0079	0.0023	95.5%	+0.0059	0.0021	96.0%	+0.0013	0.0023	96.0%	1.5002
1000	200	100.0%	-0.0003	0.0013	96.0%	+0.0007	0.0011	94.0%	+0.0033	0.0012	96.5%	3.1413

The results for the PJR estimator as shown in Table 2 revealed perfect convergence (100%) for all sample sizes in moderate coefficient conditions. As expected, the bias and MSE declined consistently with increasing sample size, indicating the estimator’s consistency. The bias of β_0 was -0.1395 , whereas MSE was 0.3273 , when $n = 10$, while the maximum MSE value was 1.0344 for β_2 as negative coefficients add extra variability in smaller sample sizes. All biases were minimal (less than

0.004), and the MSE was minimized significantly (approximately 0.001) for all parameters at $n = 1000$. The coverage probabilities did not vary much and remained constant around the 95% nominal level, ranging between 91.5% and 98%. The computation time had a linear relationship with sample sizes and varied from 0.047 to 3.141 seconds for $n = 10$ and $n = 1000$, respectively.

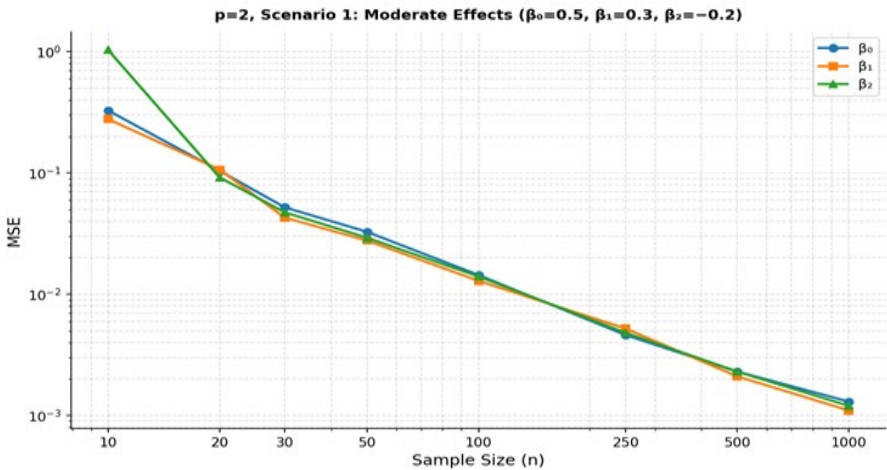


Figure 1 | MSE of all estimators vs. sample size — $p = 2$, Scenario 1 (Moderate Effects)

Figure 1 depicts moderate effects of the two predictor PJR model which shows the decreasing pattern of MSE as sample size n increases.

For large positive coefficient setups as shown in Table 3, the estimator once again converged perfectly for all sample sizes considered. In contrast to Scenario

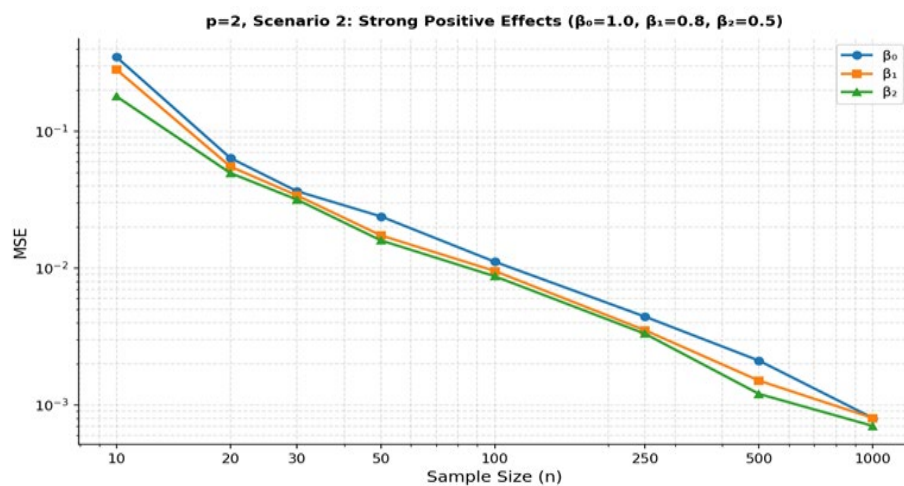
1, MSE was lower in all cases, demonstrating the advantage of increased signal strength. While at $n = 10$, the MSE of β_0 was observed to be 0.3477 , which was a marginal increase from Scenario 1, the MSE of β_1 and β_2 were noted to be 0.2818 and 0.1793 , respectively, due to their high signal strength. Biases were low and varied around zero randomly,

Table 3 | Simulation Results for Strong Positive Effects ($\beta_0 = 1.0$, $\beta_1 = 0.8$, $\beta_2 = 0.5$), $p = 2$

n	Valid	Conv%	β_0 Bias	β_0 MSE	β_0 Cov%	β_1 Bias	β_1 MSE	β_1 Cov%	β_2 Bias	β_2 MSE	β_2 Cov%	Time(s)
10	200	100.0%	-0.1327	0.3477	93.0%	-0.0584	0.2818	94.0%	-0.0158	0.1793	97.5%	0.0476
20	200	100.0%	-0.0281	0.0636	93.5%	+0.0053	0.0551	96.0%	+0.0040	0.0494	95.0%	0.0829
30	200	100.0%	-0.0572	0.0364	95.5%	+0.0008	0.0339	93.5%	-0.0054	0.0316	96.5%	0.1187
50	200	100.0%	-0.0207	0.0238	94.0%	-0.0192	0.0173	96.5%	+0.0217	0.0159	99.0%	0.1903
100	200	100.0%	-0.0228	0.0111	94.5%	-0.0025	0.0095	95.5%	-0.0045	0.0087	94.5%	0.3608
250	200	100.0%	-0.0101	0.0044	95.0%	+0.0015	0.0035	94.5%	-0.0073	0.0033	94.5%	0.9243
500	200	100.0%	-0.0076	0.0021	94.5%	-0.0018	0.0015	97.0%	-0.0025	0.0012	97.5%	1.8657
1000	200	100.0%	-0.0016	0.0008	95.5%	-0.0050	0.0008	94.0%	-0.0046	0.0007	96.5%	3.9308

implying that the estimator is almost unbiased in small samples as well. The coverage probability was good for all cases, although the interesting observation was that of β_2 at $n = 50$, with a coverage probability of 99.0%. For $n = 1000$, the MSE of all

parameters was reduced to between 0.0007–0.0008, and biases were close to zero. The high signal strength situation also yielded considerably shorter computation times relative to Scenario 1 for similar sample sizes.

**Figure 2:** MSE of all estimators vs. sample size — $p = 2$, Scenario 2 (Strong Positive Effects)

As can be seen from Figure 2, the MSEs fall sharply beyond $n=10$, signifying strong positive effects

Table 4 | Simulation Results for Mixed Large Effects ($\beta_0 = 0.2$, $\beta_1 = -0.6$, $\beta_2 = 0.9$), $p = 2$

n	Valid	Conv%	β_0 Bias	β_0 MSE	β_0 Cov%	β_1 Bias	β_1 MSE	β_1 Cov%	β_2 Bias	β_2 MSE	β_2 Cov%	Time(s)
10	200	100.0%	-0.6518	13.5966	NA%	+0.1556	7.0452	93.5%	+0.4567	12.0327	NA%	0.0533
20	200	100.0%	-0.1325	0.1739	94.5%	-0.0012	0.1264	97.5%	+0.0300	0.1446	94.0%	0.0912
30	200	100.0%	-0.0552	0.0756	96.0%	+0.0045	0.0505	95.5%	+0.0172	0.0634	96.0%	0.1448
50	200	100.0%	-0.0340	0.0459	95.0%	-0.0002	0.0299	95.5%	+0.0121	0.0318	95.5%	0.2310
100	200	100.0%	-0.0250	0.0178	97.0%	-0.0028	0.0151	93.0%	-0.0118	0.0142	96.0%	0.3991
250	200	100.0%	-0.0046	0.0078	94.0%	+0.0045	0.0051	96.5%	-0.0128	0.0057	94.5%	1.1274
500	200	100.0%	-0.0075	0.0043	94.5%	+0.0076	0.0024	96.0%	-0.0011	0.0025	95.5%	2.0680
1000	200	100.0%	-0.0036	0.0023	93.0%	+0.0066	0.0013	95.5%	-0.0014	0.0012	95.5%	4.2252

Table 4 shows that the inclusion of large mixed-sign coefficients caused very high MSE values when $n = 10$. For example, β_0 was estimated with an MSE value of 13.5966 and β_2 with 12.0327, which resulted in coverage probabilities that cannot be calculated (NA%). The rate of convergence, on the other hand, was still 100%. This means that the model did not have any problems converging, but the solutions were very unstable. Once n exceeded 20, the behaviour of the coefficients stabilized completely, as their MSEs fell to under 0.18 for each of them,

biases became negligible and random, and coverage probabilities reached acceptable levels from 93% to 98%. At $n = 1000$, all the MSEs became less than 0.0023, while biases remained smaller than 0.007 in magnitude. This example once again proves the necessity of having at least $n = 20$ when dealing with large coefficient magnitudes. This can also be seen from Figure 3, where the MSEs fall sharply beyond $n = 20$.

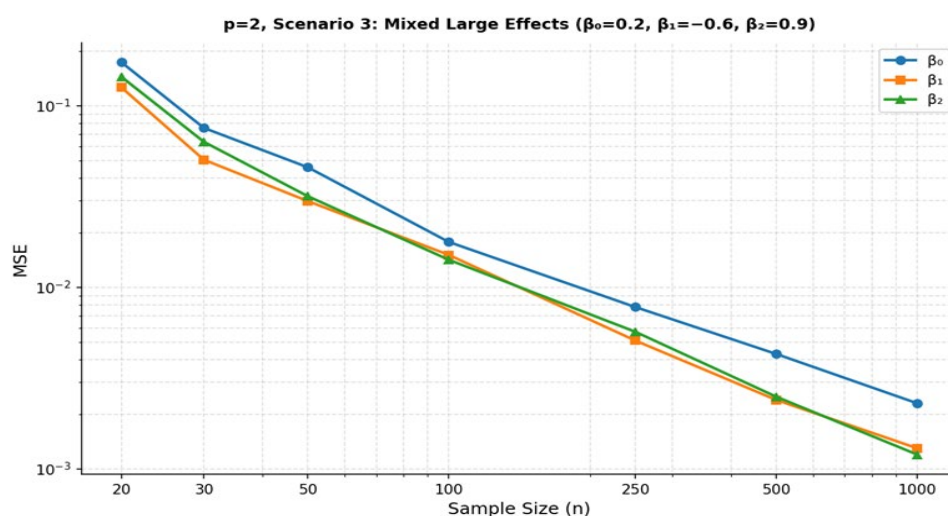


Figure 2 | MSE of all estimators vs. sample size — $p = 2$, Scenario 3 (Mixed Large Effects; $n = 10$ excluded due to extreme values)

3.3 | Application to Real Data Set

The Poisson-Juchez Regression model is applied to prior pre-term labour data set obtained from MASS: birthwt (Venables & Ripley, 2002) via a study conducted at Baystate Medical Center in 1986, to model the frequency of prior preterm births as a function of continuous maternal demographic variables. Comparison is made with the traditional

Poisson and Negative Binomial regression models to ascertain its performance.

Parameter estimates for the three count regression models considered in this study as applied to prior pre-term labour data are presented in Table 5 below for comparison.

Table 5 | Parameter Estimates (All Three Models)

Parameter	Poisson Regression				Negative Binomial Regression				Poisson-Juchez Regression			
	Estimate	Std. Error	z value	Pr(> z)	Estimate	Std. Error	z value	Pr(> z)	Estimate	Std. Error	z value	Pr(> z)
Intercept	-0.5928	1.0405	-0.5698	0.5688	-0.6870	1.1366	-0.6044	0.5455	-0.5928	1.1512	-0.5149	0.6066
Age	0.0475	0.0297	1.5984	0.1100	0.0504	0.0334	1.5109	0.1308	0.0475	0.0348	1.3624	0.1731
Lwt	-0.0175	0.0074	-2.3563	0.0185	-0.0173	0.0080	-2.1656	0.0303	-0.0175	0.0079	-2.2005	0.0278

3.4 | Model Equations and Interpretations

Table 5 presents model equations for parameter estimates of the three count regression models considered in this study and the following interpretations follow:

Poisson Regression: $\log(\mu) = -0.5928 + 0.0475 \cdot \text{age} - 0.0175 \cdot \text{lwt}$

As can be seen in Table 5, the intercept is not significant ($p = 0.569$). Maternal age shows a positive association with the log count of prior preterm labours (0.048 per year, $p = 0.110$), suggesting that older mothers tend to have a higher expected frequency, but this effect does not reach statistical significance. Pre-pregnancy weight (lwt) is a statistically significant predictor under the Poisson model (-0.018 per pound, $p = 0.019$): heavier mothers are associated with a lower expected count of prior preterm labours. The negative sign of lwt is clinically plausible, as higher pre-pregnancy weight may reflect better nutritional status and overall health.

Negative Binomial Regression: $\log(\mu) = -0.6870 + 0.0504 \cdot \text{age} - 0.0173 \cdot \text{lwt}$, $\theta = 1.0383$

With a dispersion parameter $\theta = 1.038$, the NB model introduces slightly wider standard errors compared to the Poisson. The intercept remains non-significant ($p = 0.546$). Age retains a positive coefficient (0.050 per year, $p = 0.131$) that does not reach significance. Pre-pregnancy weight (lwt) remains a significant negative predictor (-0.017 per pound, $p = 0.030$), consistent with the Poisson result. The similarity in coefficient estimates and significance patterns is expected given the mild overdispersion ($\text{Var}/\text{Mean} = 1.24$) of this outcome.

Poisson-Juchez Regression: $\log(\mu) = -0.5928 + 0.0475 \cdot \text{age} - 0.0175 \cdot \text{lwt}$, Pearson disp. = 1.1592

Coefficient values are identical to those from the Poisson model as the PJ method rescales the standard errors using the Pearson dispersion statistic (1.16) rather than re-estimating the coefficients. Consequently, the standard errors are slightly larger than under Poisson, and p-values are moderately inflated. The intercept is non-significant ($p = 0.607$). Age approaches but does not reach significance ($p = 0.173$). Notably, lwt remains a statistically significant predictor even after the dispersion correction (-0.018 per pound, $p = 0.028$), confirming that the negative association between pre-pregnancy weight and prior preterm labour count is robust across all three modelling approaches. The Pearson dispersion of 1.16 confirms that overdispersion is present but mild in this dataset.

4. | Discussion

The findings from the simulation results contained in Tables 2,3 & 4 and parameter estimates in Table 5 showed that pre-pregnancy weight (lwt) is a significant predictor and this is being upheld across all the three count models, thereby establishing it as the key determinant of prior preterm labour frequency in this study. The findings from model selection criteria revealed Poisson-Juchez regression model as the best model when compared with the traditional Poisson and negative binomial regression models, having the least AIC and BIC as can be seen from Table 6.

5. | Conclusion

The newly proposed count regression model (Poisson-Juchez regression model) presents a significant advancement in modeling over-dispersed count data. Through simulation, parameter estimation and comprehensive analysis, we have

demonstrated that the Poisson-Juchez regression model performed better than the traditional Poisson and Negative Binomial regression models, as its application to prior pre-term labour had proven its practical flexibility and robustness. The goodness-of-fit tests conducted had also indicated that the Poisson-Juchez regression model provides a superior fit, as evidenced by better AIC and BIC scores. The study is, however, limited to count data

(non-negative integer-value) modelling. This study could form a basis for future research in proposing regression models for hybrid distributions. The study can be extended by considering a situation where the model dimension p is not equal to two predictor variables, that is, when model dimension p is greater than two.

References

- Altun, E. (2018). A new zero-inflated regression model with application. *Journal of Statisticians: Statistics and Actuarial Sciences*, 11(2), 73–80.
- Altun, E. (2019). A new model for over-dispersed count data: Poisson quasi-Lindley regression model. *Mathematical Sciences*, 13, 241–247. <https://doi.org/10.1007/s40096-019-0293-5>
- Altun, E. (2020). A new one-parameter discrete distribution with associated regression and integer-valued autoregressive models. *Mathematica Slovaca*, 70(4).
- Blackburn, M. L. (2020). The relative performance of Poisson and negative binomial regression estimators. *Oxford Bulletin of Economics and Statistics*, 77, 605–616.
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24, 127–135.
- Breslow, N. (1990). Tests of hypotheses in over dispersed Poisson regression and other quasi-likelihood models. *Journal of the American Statistical Association*, 85, 565–571.
- Breslow, N., & Clayton, D. (1993). Approximate inference in generalized linear mixed models *Journal of the American Statistical Association*, 88(421), 9–25.
- Cameron, A. C., & Trivedi, P. K. (1990). Regression-based tests for overdispersion in the Poisson model. *Journal of Econometrics*, 46(3), 347–364.
- Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (2nd ed.). Cambridge University Press.
- Deb, P., & Trivedi, P. K. (1997). Demand for medical care by the elderly: A finite mixture approach. *Journal of Applied Econometrics*, 12, 313–336.
- Denuit, M., Maréchal, X., Pitrebois, S., & Walhin, J. (2007). *Actuarial modelling of claim counts: Risk classification, credibility and bonus-malus systems*. John Wiley & Sons.
- Dupuy, J.-F. (2018). *Statistical methods for overdispersed count data*. Elsevier.
- Fávero, L. P. (2019). Count data regression analysis: Concepts, overdispersion detection, zero-inflation identification, and applications with R. *Practical Assessment, Research & Evaluation*, 26(13). <https://scholarworks.umass.edu/pare/vol26/iss1/13/>

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Klakattawi, H. S., Vinciotti, V., & Yu, K. (2018). A simple and adaptive dispersion regression model for count data. *Entropy*, 20(2), 142.

Koyejo, S. O., Akomolafe, A. A., Oseni, B. M., & Ojo, O. O. (2026). A new Poisson-Juchez distribution with its properties and application. *Pacific Journal of Science and Technology*, 27(1), 73–87. <https://www.akamai.university/pacific-journal-of-science-and-technology.html>

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment. *Transportation Research Part A: Policy and Practice*.

McCulloch, C. E., Searle, S. R., & Neuhaus, J. M. (2011). *Generalized, linear, and mixed models* (2nd ed.). John Wiley & Sons. <https://www.perlego.com/book/1006716/generalized-linear-and-mixed-models>

Payne, J. R., Dise, N. B., Field, C. D., Dore, A. J., Caporn, S. J. M., & Stevens, C. J. (2017). Nitrogen deposition and plant biodiversity: Past, present and future. *Frontiers in Ecology and the Environment*, 15(8), 431–436.

Payne, E. H., Gebregziabher, M., Hardin, J. W., Ramakrishnan, V., & Egede, L. E. (2018). An empirical approach to determine a threshold for assessing overdispersion in Poisson and negative binomial models for count data. *Communications in Statistics - Simulation and Computation*, 47(6), 1722–1738.

Radhakumari, M., Muhammed, R. I., Damodaran, S., Christophe, C., Soman, L. N., & Damodaran, S. S. (2022). On discrete Poisson-Mirra distribution: Regression, INAR(1) process. *Axioms*, 11, 193. <https://doi.org/10.3390/axioms11050193>

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer. <https://doi.org/10.1007/978-0-387-21706-2>

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software*, 27(8), 1–25. <https://doi.org/10.18637/jss>

Zeviani, W. M., Ribeiro Jr., P. J., Shimakura, S. E., & Muniz, J. A. (2014). The Gamma-count distribution in the analysis of experimental underdispersed data. *Journal of Applied Statistics*, 41(12), 2616–2626. <https://doi.org/10.1080/02664763.2014.922168>