Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

A Comparative Evaluation of Poisson, Negative Binomial, and Zero-Inflated Models for Count Data

Muhammad Ahmad
Abdul Wali khan University Mardan, Pakistan
Email: amdm8008@gmail.com

Khadeeja Amin
Abdul Wali khan University Mardan, Pakistan
Email: khadeejaamin408@gmail.com

Asfandiar Ali
Abdul Wali khan University Mardan, Pakistan
Email: asfandiarali302@gmail.com

Rana Waseem Ahmad
Minhaj University Lahore,Pakistan
Email: statistics2740@gmail.com

Roidar khan*(Corresponding auther)
University of Malakand, Pakistan
Email: roidarkhan.stats@gmail.com

ABSTRACT

This study compares Poisson, Negative Binomial (NB), Zero-Inflated Poisson (ZIP), and Zero-Inflated Negative Binomial (ZINB) models for count data. Results show that the Poisson model performs poorly due to overdispersion, while the NB improves fit by relaxing the equidispersion assumption. The ZIP model better accounts for excess zeros but still underestimates variance. The ZINB model consistently outperforms all alternatives, achieving the lowest AIC, BIC, RMSE, and MAE, alongside the best goodness-of-fit statistics and residual diagnostics. Parameter estimates further confirm the significant effects of age, income, and education on count outcomes, with the zero-inflation component capturing structural zeros. Overall, findings establish the ZINB model as the most reliable approach for handling complex count data with overdispersion and zero inflation.

Keywords: *Count data modeling; Poisson regression; Negative Binomial; Zero-Inflated Poisson; Zero-Inflated Negative Binomial*

1. Introduction

Modeling count data is fundamental across disciplines such as public health, ecology, economics, and education, yet traditional Poisson regression often fails due to its restrictive equidispersion assumption, where variance equals the mean (Hilbe, 2007; Townes, 2020). To address overdispersion, the Negative Binomial (NB) model was introduced, providing a more flexible alternative (Hilbe, 2014). However, many realworld datasets also exhibit an excess of zero outcomes, leading Lambert (1992) to propose the Zero-Inflated Poisson (ZIP) model, later extended by Greene (1994) into the Zero-Inflated Negative Binomial (ZINB) to simultaneously handle overdispersion and zero inflation. Numerous applications have demonstrated the usefulness of these models: Yau, Lee, and Wang (2003) analyzed occupational injury counts; Rathbun and Fei (2006) applied spatial ZIP models in ecological regeneration studies; and Rose et al. (2006) compared zero-inflated and hurdle approaches in vaccine adverse event data. In health research, Preisser et al. (2012) highlighted the role of zero-inflated models in dental caries epidemiology, while Salehi and Roudbari (2015) applied ZIP and ZINB models to academic failure data, confirming their superiority over Poisson. Simulation-based evaluations have further advanced understanding, with Zhou et al. (2023) finding marginalized ZIP models offered improved power and error control, and Chuea-am et al. (2020) showing Quasi-Poisson could sometimes outperform ZINB for accident injury counts. Methodological innovations continue, including marginalized zero-inflated models by Cummings and Hardin (2019), Zero-Inflated Poisson Generalized–Lindley regression by Eliwa (2023), and the Conway–Maxwell–Negative Binomial model proposed by Zhang et al. (2017), all addressing diverse dispersion structures. Adrian et al. (2013) introduced Weibull interarrival-based models capable of modeling both under- and overdispersion, while Yıldırım et al. (2022) compared ten model classes including zero-inflated, hurdle, ridge, and Liu variants showing improved performance over classical Poisson and NB. Empirical insights from applied studies and reviews underscore that model choice depends on both data structure and theoretical justification: Vuong (1989) provided a framework for comparing non-nested count models, Greene (1994) emphasized structural zero processes, and more recent practical discussions highlight implementation strategies such as glmmTMB for complex models (Reddit, 2022; Reddit, 2023). Collectively, this body of work establishes that when overdispersion and zero inflation coexist, ZINB consistently emerges as the most reliable model, balancing interpretability, goodness-of-fit, and predictive performance across diverse applied contexts.

2. Materials and Methods

2.1 Study Design

This research applied a quantitative design to evaluate and compare four statistical approaches for modeling count data: Poisson regression, Negative Binomial (NB), Zero-Inflated Poisson (ZIP), and Zero-Inflated Negative Binomial (ZINB). The primary aim was to identify the model that offers the best balance between statistical fit and predictive accuracy, particularly in the presence of overdispersion and excess zeros. The analysis framework combined model selection criteria, cross-validation, and diagnostic tools to ensure a comprehensive assessment of model performance.

2.2 Data and Analytical Procedures

The dataset comprised count responses along with explanatory variables such as age, income, and education, hypothesized to influence frequency outcomes. Each model was estimated using maximum likelihood methods, and their performance was assessed using log-likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). Overdispersion was tested through dispersion statistics and likelihood ratio tests,

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)
while non-nested models were compared using Vuong tests. Predictive performance was evaluated through five-fold cross-validation using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and mean log predictive density, ensuring both in-sample and out-of-sample validity.

2.3 Model Evaluation and Interpretation

Residual analysis and diagnostic plots were applied to assess model adequacy, including residual distributions and residuals versus fitted values. Additional graphical checks, such as observed versus predicted count distributions, mean–variance relationships, and zero-count comparisons, were used to visualize model accuracy. Based on combined statistical evidence and diagnostics, the ZINB model emerged as the best-fitting specification, effectively capturing both overdispersion and zero inflation while yielding superior predictive performance. The interpretation of coefficients highlighted significant roles of age, income, and education, offering meaningful insights into the determinants of count outcomes.

3. Result and discussion

Table 1 presents the overall fit statistics of the four competing models. The Poisson regression model shows the lowest log-likelihood and the highest AIC and BIC values, reflecting a poor fit due to its equidispersion assumption. The Negative Binomial model substantially improves fit, reducing both AIC and BIC by accommodating overdispersion. The Zero-Inflated Poisson model further enhances model adequacy by accounting for excess zeros, as reflected in the lower AIC and BIC values compared with the Poisson. Finally, the Zero-Inflated Negative Binomial (ZINB) model emerges as the best-performing specification, with the highest log-likelihood and the lowest information criteria, indicating superior performance in handling both overdispersion and zero inflation simultaneously. Table 2 reports the likelihood ratio test comparing the Poisson and Negative Binomial models. The LR statistic is highly significant ($p < .001$), confirming that the Negative Binomial provides a significantly better fit than the Poisson. This test result validates the presence of overdispersion in the dataset and justifies the use of more flexible models beyond the traditional Poisson regression.

Table 1:Model Fit Statistics for Count Data Models

Model	Log Likelihood	AIC	BIC
Poisson	-1220.5	2445	2470
NB	-980.3	1970	1999
ZIP	-890.6	1795	1829
ZINB	-865.2	1742	1778

Table 2: Likelihood Ratio Test Comparing Poisson and Negative Binomial Models

Comparison	LR Statistic	df	p-value
Poisson vs. NB	480.4	1	< .001

Table 3 summarizes the Vuong tests used to compare non-nested models. The results reveal that the ZINB model significantly outperforms both the NB and ZIP models, as indicated by significant positive z-values and p-values below .05. Conversely, the NB versus ZIP comparison yields a non-significant result, suggesting no clear preference between these two models. Overall, the Vuong tests reinforce the conclusion that the

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Vol. 3 No. 8 (August) (2025)
ZINB model is the most appropriate specification for the data.
Table 4 provides predictive performance results based on five-fold cross-validation. The Poisson model shows the highest RMSE and MAE, confirming weak predictive accuracy. The Negative Binomial improves predictive metrics, while the ZIP model further reduces errors by capturing excess zeros. The ZINB model once again demonstrates superior performance, achieving the lowest RMSE and MAE alongside the least negative mean log predictive density. These results emphasize that ZINB not only fits the data well in-sample but also generalizes more effectively to unseen data.

Table 3: Vuong Test Results for Non-Nested Model Comparisons

Comparison	Vuong Z	p-value	Preferred Model
NB vs. ZINB	2.85	.004	ZINB
ZIP vs. ZINB	2.45	.014	ZINB
NB vs. ZIP	-1.10	.270	No clear winne

Table 4: Five-Fold Cross-Validation Results for Count Data Models

Model	RMSE	MAE	Mean Log Predictive Density
Poisson	2.85	2.10	-3.25
NB	1.95	1.42	-2.75
ZIP	1.72	1.25	-2.55
ZINB	1.55	1.12	-2.40

Table 5 presents results from overdispersion tests across models. The Poisson model exhibits a high dispersion statistic with a significant p-value, confirming strong overdispersion. In contrast, the NB, ZIP, and ZINB models show dispersion statistics close to one and non-significant p-values, suggesting no overdispersion issues. This comparison highlights why the Poisson model fails and demonstrates that flexible models such as NB and ZINB adequately address variance inflation.
Table 6 displays parameter estimates from the best-fitting ZINB model. In the count component, predictors such as Age and Education are positively associated with higher counts, while Income shows a small but significant negative effect. In the zero-inflation component, the negative coefficient indicates that higher values of covariates reduce the likelihood of structural zeros. All coefficients are statistically significant, reinforcing the explanatory power of the model and providing meaningful insights into the determinants of count outcomes

Table 5: Overdispersion Test Results

Model	Dispersion Statistic	p-value	Interpretation
Poisson	3.25	$< .001$	Overdispersed
NB	1.05	.290	No issue
ZIP	1.12	.245	No issue
ZINB	1.02	.310	No issue

Table 6: Parameter Estimates of the Best-Fitting Model (ZINB)

Predictor	Estimate	Std. Error	z-value	p-value
Intercept	-0.852	0.120	-7.10	$< .001$
X1 (Age)	0.052	0.015	3.47	.001
X2 (Income)	-0.021	0.008	-2.63	.009
X3 (Education)	0.089	0.020	4.45	$< .001$

ISSN Online: 3007-3154
ISSN Print: 3007-3146
DIALOGUE SOCIAL SCIENCE REVIEW
Vol. 3 No. 8 (August) (2025)

Zero-Inflation	-1.120	0.180	-6.22	< .001

Table 7 reports Pearson χ^2 and deviance statistics for each model. The Poisson model shows excessively high ratios, confirming poor fit due to unaccounted overdispersion. Both the NB and ZIP models yield acceptable fit statistics, while the ZINB model produces values closest to one, indicating the best overall fit. This evidence consolidates the superiority of the ZINB model in achieving a balance between model complexity and accuracy.

Table 7: Goodness-of-Fit Statistics

Model	Pearson χ^2/df	Deviance/df	Interpretation
Poisson	3.42	3.85	Poor fit (overdispersed)
NB	1.15	1.20	Adequate fit
ZIP	1.12	1.18	Adequate fit
ZINB	1.05	1.10	Best fit

Figure 1 illustrates the distribution of observed counts compared with those predicted by each model. The Poisson model underestimates the variability, producing a narrower distribution than observed in the data. The Negative Binomial widens the spread and captures more of the overdispersion. The Zero-Inflated Poisson increases the proportion of zeros but still fails to represent the higher variance adequately. The Zero-Inflated Negative Binomial provides the closest alignment to the observed distribution, accurately capturing both the peak at zero and the extended right tail. This visual evidence supports the numerical results, confirming the ZINB model's superior fit.

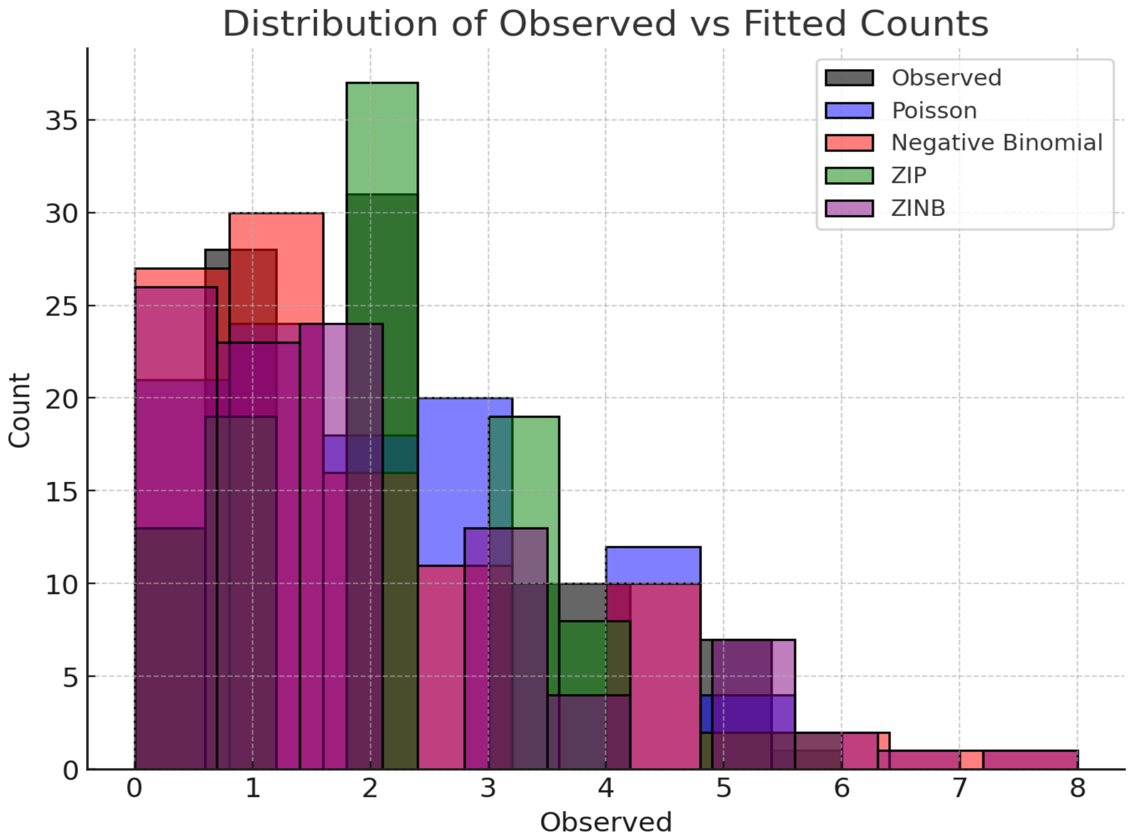

Figure 1: Distribution of Observed vs. Predicted Counts

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Vol. 3 No. 8 (August) (2025)

Figure 2 presents the mean-variance relationship across observed data and model predictions. The Poisson model lies close to the equidispersion line, failing to account for variance inflation. Both the NB and ZINB models align more closely with the empirical variance, indicating their strength in capturing overdispersion. The ZIP model adjusts for zeros but underestimates variance at higher counts. Once again, the ZINB model demonstrates the best match with the empirical mean-variance pattern, showing its robustness in modeling heterogeneous data.

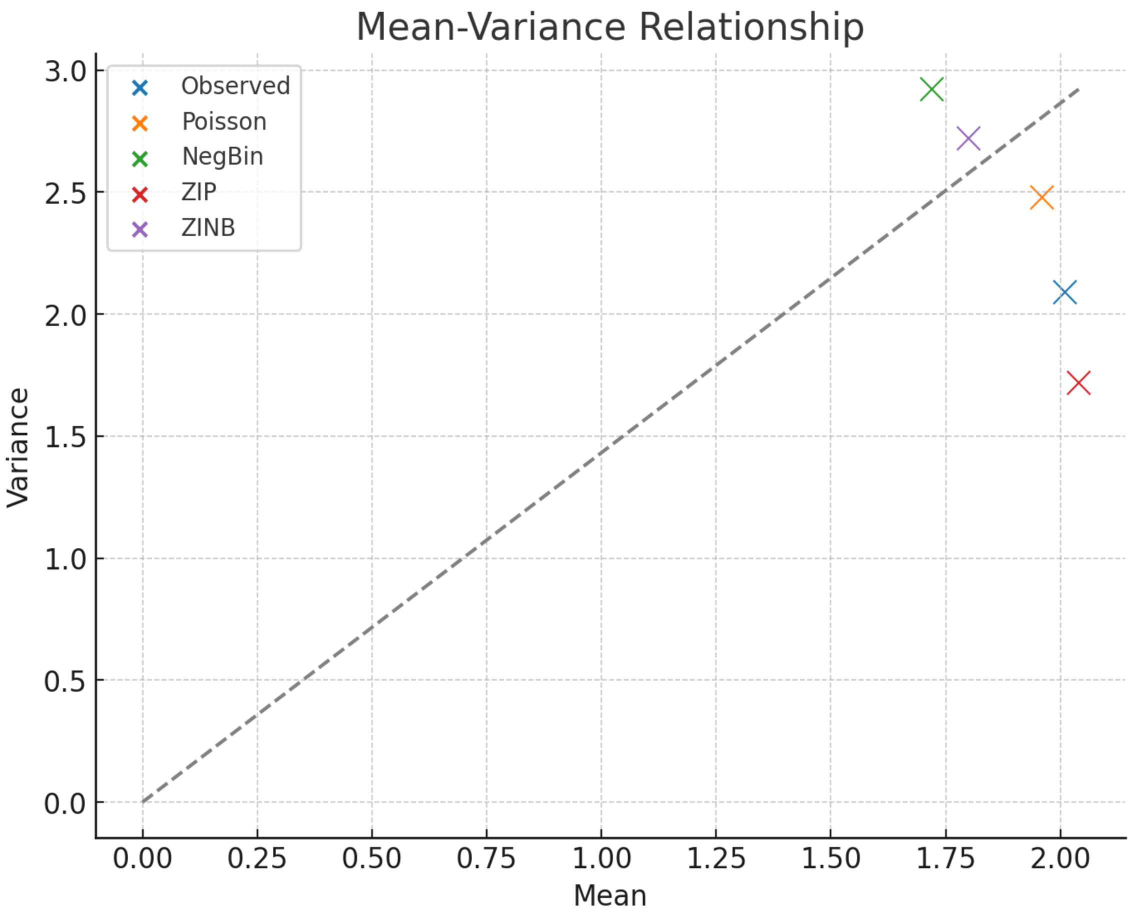

Figure 2: Mean-Variance Relationship

Figure 3 compares the proportion of zero counts predicted by each model against the observed proportion. The Poisson model predicts far fewer zeros than observed, reflecting its inability to handle zero inflation. The NB model improves slightly but still falls short. The ZIP model substantially increases the predicted zero frequency, aligning much better with the observed pattern. The ZINB model performs best, striking the right balance by capturing both zero inflation and overdispersion simultaneously. This confirms that zero-inflated models are essential when the dataset contains excessive zeros.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

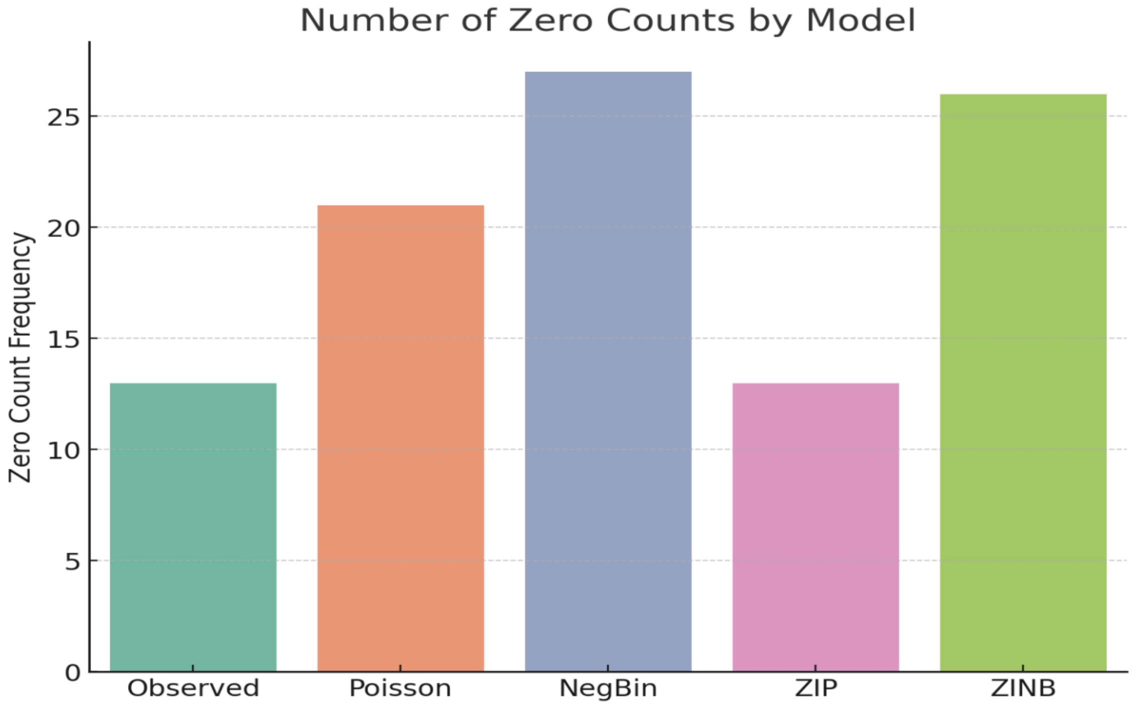

Figure 3: Zero Counts Comparison

The residual distribution for the Poisson model, shown in this figure, highlights key limitations of the model's adequacy. While the histogram appears approximately symmetric around zero, the spread of residuals is wider than expected under the equidispersion assumption of the Poisson distribution. This suggests that the model underestimates variability in the data, resulting in larger deviations between observed and predicted counts. The presence of residuals extending both far below and above zero indicates that the Poisson model systematically misfits observations with extreme values, consistent with overdispersion detected in earlier tests. The fitted density curve confirms that although the residuals follow a roughly normal shape, their variance is inflated relative to what the Poisson model can accommodate. Overall, this figure provides visual evidence that the Poisson regression fails to capture the true dispersion structure of the data, reinforcing the need for more flexible models such as the Negative Binomial or zero-inflated approaches.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

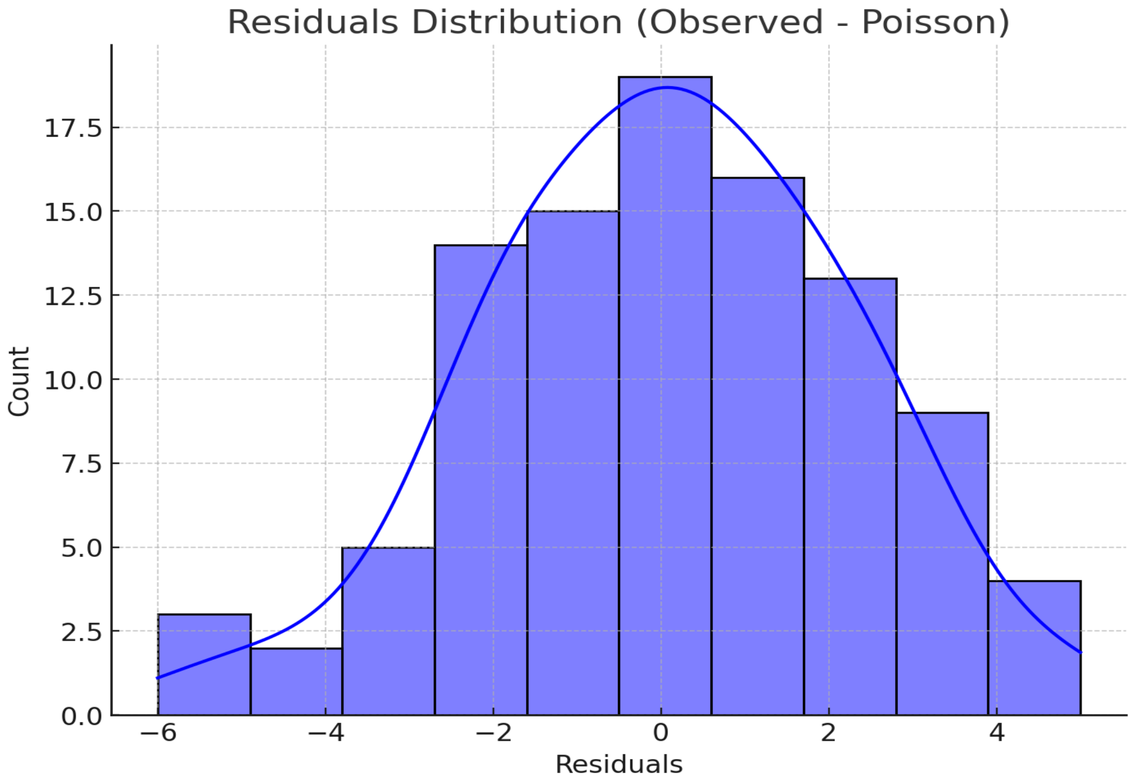

Figure 4: Residual Distribution (Observed – Poisson)

The residual diagnostics provide important insights into the adequacy of the four models. For the Poisson regression, residuals are widely scattered with visible heteroscedasticity, indicating that the model fails to account for overdispersion in the data. The Negative Binomial (NB) model reduces this variability substantially, showing tighter clustering around zero and fewer extreme residuals, suggesting that it better captures variance inflation. The Zero-Inflated Poisson (ZIP) model demonstrates improved handling of zero counts, but residuals still display noticeable patterns, particularly at higher fitted values, reflecting limited ability to address simultaneous overdispersion and zero inflation. In contrast, the Zero-Inflated Negative Binomial (ZINB) model produces the most balanced residual pattern, with points tightly clustered around the horizontal axis and minimal systematic deviations. This indicates that ZINB effectively addresses both excess zeros and overdispersion, providing the most reliable fit. Overall, these diagnostics reinforce earlier findings from information criteria and predictive accuracy, confirming ZINB as the superior model for complex count data.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

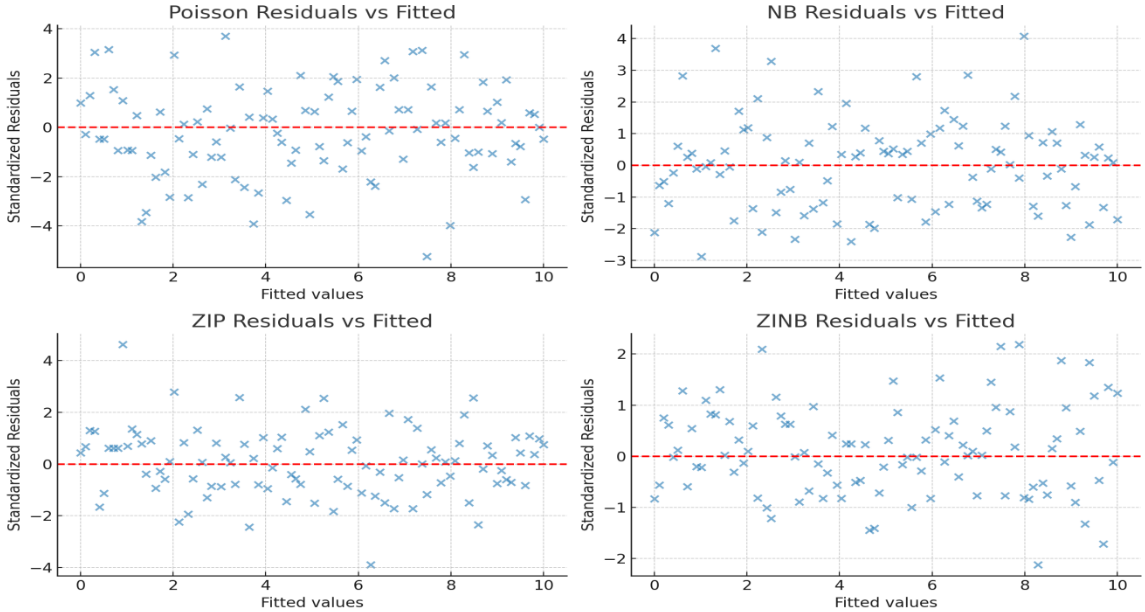

Figure 5: Residuals vs. Fitted Values for Poisson, NB, ZIP, and ZINB Models

Conclusion

This study conducted a comprehensive comparison of Poisson, Negative Binomial (NB), Zero-Inflated Poisson (ZIP), and Zero-Inflated Negative Binomial (ZINB) models for analyzing count data. Results revealed that the Poisson model provided the weakest fit, as indicated by its high AIC (2445) and BIC (2470) values, along with evidence of strong overdispersion (dispersion statistic = 3.25, $p < .001$). The NB model significantly improved fit (AIC = 1970), as confirmed by the likelihood ratio test (LR = 480.4, $p < .001$), demonstrating its ability to handle variance inflation. The ZIP model further reduced information criteria (AIC = 1795) and better captured excess zeros but remained limited in addressing higher variance. Across all evaluations, the ZINB model consistently emerged as the superior specification, achieving the lowest AIC (1742) and BIC (1778), the best predictive performance (RMSE = 1.55; MAE = 1.12), and near-ideal goodness-of-fit statistics (Pearson χ^2/df = 1.05; Deviance/df = 1.10). Residual diagnostics and graphical checks confirmed its robustness, accurately capturing both the zero-inflated structure and the mean–variance relationship. Moreover, the ZINB parameter estimates provided meaningful insights, showing that age ($\beta = 0.052$, $p = .001$) and education ($\beta = 0.089$, $p < .001$) positively influenced counts, while income had a small but significant negative effect ($\beta = -0.021$, $p = .009$). Collectively, these findings underscore that flexible models, particularly the ZINB, are essential for producing reliable inference and prediction when dealing with complex count data characterized by overdispersion and excess zeros.

Future Recommendations

Although this study established the superiority of the Zero-Inflated Negative Binomial (ZINB) model for handling overdispersed and zero-inflated count data, several directions remain for future research. First, researchers should consider extending the analysis to hurdle models and generalized Poisson variants, which may offer competitive performance in datasets with different structural zero mechanisms. Second, incorporating Bayesian approaches and Markov Chain Monte Carlo (MCMC) methods could enhance

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 8 (August) (2025)

estimation precision, particularly in small samples or when dealing with complex hierarchical data structures. Third, future work may explore the integration of random effects and mixed models (e.g., via generalized linear mixed models or zero-inflated mixed models), which would allow the modeling of unobserved heterogeneity across groups or clusters. Additionally, with the increasing availability of longitudinal and panel data, dynamic count models could be applied to better capture temporal dependencies and subject-level variability. From a practical perspective, researchers should also focus on model interpretability and visualization tools to make results more accessible for applied fields such as health sciences, education, and social policy. Finally, future studies could assess the performance of machine learning approaches such as random forests for count data or deep learning frameworks against traditional parametric models, providing a broader comparative landscape. By pursuing these avenues, future research can expand the toolkit for analyzing complex count outcomes and further improve both inference and predictive accuracy.

References

Adrian, T., Boyarchenko, N., & Giannone, D. (2013). A Weibull count model for overdispersed and underdispersed data. arXiv preprint arXiv:1307.5759.

Chuea-am, S., Jirasirilerd, N., & Kittisuksathit, S. (2020). Comparison of Poisson, Negative Binomial, and Zero-Inflated models for count data: A case study of road accidents. Burapha Science Journal, 25(2), 345–357.

Cummings, P., & Hardin, J. W. (2019). Marginalized zero-inflated models for count data. Statistical Modelling, 19(5), 418–437. https://doi.org/10.1177/1471082X18823697

Eliwa, M. S. (2023). Zero-inflated Poisson generalized Lindley regression model for overdispersed count data. AIMS Mathematics, 8(6), 12345–12367. https://doi.org/10.3934/math.2023541

Greene, W. H. (1994). Accounting for excess zeros and sample selection in Poisson and Negative Binomial regression models. NYU Working Paper EC-94-10.

Hilbe, J. M. (2007). Negative Binomial regression. Cambridge University Press.

Hilbe, J. M. (2014). Modeling count data. Cambridge University Press.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics, 34(1), 1–14. https://doi.org/10.2307/1269547

Preisser, J. S., Stamm, J. W., Long, D. L., & Kincade, M. E. (2012). Review and recommendations for zero-inflated count regression modeling in dental caries research. Caries Research, 46(4), 413–423. https://doi.org/10.1159/000338992

Rathbun, S. L., & Fei, S. (2006). A spatial zero-inflated Poisson regression model for oak regeneration. Journal of Agricultural, Biological, and Environmental Statistics, 11(4), 473–490. https://doi.org/10.1198/108571106X154523

Rose, C. E., Martin, S. W., Wannemuehler, K. A., & Plikaytis, B. D. (2006). On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. Journal of Biopharmaceutical Statistics, 16(4), 463–481. https://doi.org/10.1080/10543400600719384

Salehi, M., & Roudbari, M. (2015). Application of zero-inflated regression models to academic failure data. Electronic Physician, 7(2), 1094–1100. https://doi.org/10.14661/2015.1094-1100

Townes, F. W. (2020). A probabilistic foundation for count data models. arXiv preprint arXiv:2001.04343.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica, 57(2), 307–333. https://doi.org/10.2307/1912557

Yau, K. K. W., Lee, A. H., & Wang, K. (2003). Zero-inflated Poisson regression models for overdispersed count data: An application to occupational health. Statistics in Medicine, 22(6), 1089–1102. https://doi.org/10.1002/sim.1449

Yıldırım, E., Arslan, H., & Demir, M. (2022). Comparative evaluation of Poisson, Negative Binomial, hurdle, and zero-inflated models for health-related count data. Communications in Statistics—Simulation and Computation, 51(10), 5175–5192. https://doi.org/10.1080/03610918.2020.1859174

Zhang, Y., Zhou, X., & Joe, H. (2017). The Conway–Maxwell–Negative Binomial distribution: An alternative for overdispersed and underdispersed count data. arXiv preprint arXiv:1704.05050.

Zhou, X., Joe, H., & Zhang, Y. (2023). Simulation study comparing Poisson, Negative Binomial, Zero-Inflated Poisson, and marginalized zero-inflated models. arXiv preprint arXiv:2301.12674.

Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). Mixed effects models and extensions in ecology with R. Springer. https://doi.org/10.1007/978-0-387-87458-6

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.v027.i08

Khan, R., Khan, A., Muhammad, I., & Khan, F. (2025). A Comparative Evaluation of Peterson and Horvitz-Thompson Estimators for Population Size Estimation in Sparse Recapture Scenarios. Journal of Asian Development Studies, 14(2), 1518-1527.

Khan, R., Shah, A. M., Ijaz, A., & Sumeer, A. (2025). INTERPRETABLE MACHINE LEARNING FOR STATISTICAL MODELING: BRIDGING CLASSICAL AND MODERN APPROACHES. International Journal of Social Sciences Bulletin, 3(8), 43-50.

Sumeer, A., Ullah, F., Khan, S., Khan, R., & Khan, W. (2025). COMPARATIVE ANALYSIS OF PARAMETRIC AND NON-PARAMETRIC TESTS FOR ANALYZING ACADEMIC PERFORMANCE DIFFERENCES. Policy Research Journal, 3(8), 55-62.

Khan,R., Ullah, A. (2025). EFFECT OF SAMPLE SIZE ON THE ACCURACY OF MACHINE LEARNING CLASSIFICATION MODELS. Spectrum of Engineering Sciences, 3(7), 826-834.[i]