Current Science
CS 5 (3), 3059-3071 (2025)
Received 20 March 2025 | **Revised** 26 April 2025 | **Accepted** 28 May. 2025 | **Online Available** 16 June 2025
https://doi.org/ 10.52845/CS/2025-5-3-75

Current Science
Current Science

OPEN ACCESS **ISSN(O)2795-8639**

Original Article

Analysis of Ammunition Malfunctions Based on the Negative Binomial-Lindley Distribution

Cenyu Hu[1], Ling Fang[2], Xianming Shi[1*], Xiaolu Cui[3], Yalong Wang[1]

[1]Department of Equipment Command and Management，Shijiazhuang Campus of Army Engineering University，Shijiazhuang, China

[2]Army Logistics Academy, Chongqing, China

[3]Command and Control Engineering College，Army Engineering University of PLA, Nanjing, China

[5]Department of Equipment Command and Management，Shijiazhuang Campus of Army Engineering University，Shijiazhuang, China

[*]Corresponding Author: Xianming Shi

Abstract:
The reliability of ammunition systems is of vital importance to the combat effectiveness and safety in modern warfare. However, the failure data often exhibit significant characteristics of excessive zeros and excessive dispersion, which cannot be effectively fitted by traditional data. This paper aims to establish a new count data model - the NB-L model, to more accurately capture the complex features of ammunition failure data. Under the framework of Bayesian hierarchical models, parameter estimation is conducted using the Markov Chain Monte Carlo (MCMC) method, and Bayesian methods are adopted to integrate prior information and quantify uncertainty. After conducting an empirical analysis on a typical ammunition production quality control dataset with zero inflation and excessive dispersion, a comparative analysis is made between the Bayesian Poisson distribution and the negative binomial distribution models to verify the accuracy of the model. Based on the Deviance Information Criterion (DIC) and Deviance evaluation. The NB-L GLM exhibits the optimal performance in fitting ammunition failure data with covariates. The results show that the NB-L distribution provides a powerful alternative tool for such complex count data and offers more precise statistical basis for related decisionsexpressions
Keywords: ammunition systems , negative binomial-Lindley distribution , bayesian inference, regression analysis

1. Introduction

The reliability of ammunition systems is the foundation for the sustained combat capability of equipment[1]. Under modern combat conditions, the stability of ammunition performance can directly affect combat effectiveness and personnel safety[2].

Owing to the heterogeneous and intricate characteristics of the battlefield environment, the reliability of the ammunition system faces significant challenges. Statistics show that environmental factors contribute to over 30% of ammunition malfunctions in actual combat scenario. Therefore, establishing a statistical model that can precisely quantify the influence of environmental conditions and design parameters on ammunition malfunctions is not only a theoretical requirement in equipment research and development but also an important practical foundation for enhancing the efficiency of battlefield logistics support and the accuracy of tactical decision-making[3-5].

Among the data of ammunition malfunctions, the ammunition system is mostly in a fault-free condition (i.e., the number of malfunctions is zero). In low-temperature environments, the physical properties of materials may significantly reduce the probability of malfunctions, resulting in an abnormally high proportion of zero values in the data[6]. However, in extreme or harsh environments (such as high temperature, high humidity, or strong electromagnetic interference), the ammunition system may malfunction, and the non-zero malfunction counts exhibit an excessive discreteness characteristic. This makes traditional technical models difficult to capture the statistical features of such data due to the limitations of their distribution assumptions[7].

The traditional fault analysis methods mostly employ the Poisson distribution, but the Poisson distribution cannot effectively capture discrete phenomena due to its characteristic that the variance is equal to the mean. For the standard negative binomial distribution, although introducing discrete parameters can alleviate some overdispersion problems, due to the difficulty of the single gamma mixture structure in adapting to data with excessive zeros, the accuracy is limited in extreme environments[8]. In recent years, the mixture distribution model has become a hot topic for solving such problems due to its flexibility and adaptability[9-10]. By introducing multi-level structures or additional parameters, it can more accurately describe the potential characteristics of complex data. The Lindley distribution has attracted much attention for its asymmetry in the positive real number domain and its fitting for zero-valued data[11]. The extended model based on the Lindley distribution has been verified in various fields such as environment, economy, and transportation to have strong explanatory power for complex data[12]. Zamani and Ismail (2010) were the first to propose the negative binomial -Lindley distribution, combining the negative binomial model with the Lindley distribution to achieve methodological progress. Lord and Geedipally used the NB-L distribution to study collision frequency datasets, which have a large number of zero-overflows[13]. Their comparative analysis shows that in the face of excessive zero observations, the NB-L distribution consistently outperforms the standard negative binomial and Poisson distributions in terms of prediction accuracy and goodness-of-fit indicators[14]. Subsequently, Geedipally et al. (2012) expanded on this and proposed a generalized linear model (GLM) implementation of the NB-L distribution.

This paper mainly adopts the Bayesian method to construct the hierarchical structure of the negative binomial-Lindley model within the framework of Generalized Linear Model (GLM) in order to achieve parameter estimation[15].Bayesian inference has unique advantages in integrating prior information, especially being applicable to the common small sample situations in ammunition failure analysis, and it realizes comprehensive uncertainty quantification through the posterior distribution[16].

Furthermore, by elaborating on the theoretical basis of NB-L, and applying this model to ammunition failure data for empirical analysis, our aim is to evaluate its parameters and compare them with existing count models. Through this analysis, we aim to demonstrate the superiority of this model in terms of adaptability and practicality, thereby enhancing the feasibility of researchers when dealing with complex data[18-19].

NB-Lindley Distribution Model

2.1 NB Distribution

Before constructing the NB-L model, it is necessary to establish the theoretical basis of the negative binomial distribution[3-4]. Based on mathematical derivation of the limit form parameterization of a series of independent Bernoulli trials, the PMF can be expressed as:

- $$P(Y=y;r,p)=\frac{\Gamma(y+r)}{\Gamma(y+1)\Gamma(r)}(1-p)^{y}(p)^{r} \qquad (1)$$

Due to the excessive number of zero and the significant heterogeneity of the data, we adopt the reparameterization strategy to represent the original parameter p by the average number of failures μ and the discrete parameter r :

$$p = \frac{\mu}{\mu + r} \qquad (2)$$

The result can be further derived from Formula (2),

$$E(Y) = \mu \text{ and } Var(Y) = \mu + \frac{\mu^2}{r} \qquad (3)$$

2.2 Lindley Distribution

The NB model may still exhibit potential limitations or failure rates[6-7].Variations in ammunition batches, production costs, and complex operational capabilities contribute to differences in inherent failure propensity or occurrence rates λ.For the modeling of the underlying heterogeneity of this potential failure rate λ, we introduce the Lindley distribution which can be interpreted as a mixture of $E(\theta)$ and $gamma(2,\theta)$.

Let the potential failure rate λ be a random variable that follows the Lindley distribution and the PMF of its is

$$f(\lambda;\theta) = \frac{(1+\lambda)\theta^2}{1+\theta} e^{-\theta\lambda} \qquad (4)$$

where λ represents the potential failure rate, which quantifies the inherent failure risk level of the unit; θ is the tail behavior and density near the zero value of the Lindley distribution, which determines the distribution characteristics of the failure rate λ.

Figure 1 has a unique heavy-tailed characteristic compared with the traditional exponential distribution[9].

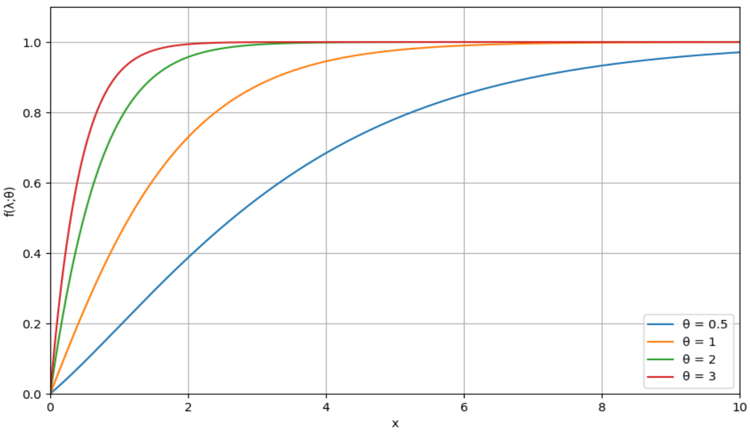

Figure 1 The pmf plots of θ with the specified value of parameters

The MGF of λ can be obtained by calculation as follows,

$$M_\lambda(t;\theta) = E[e^{t\lambda}] = \frac{\theta^2}{(1+\theta)(\theta-t)}(1+\frac{1}{\theta-t}); t>0 \qquad (5)$$

As well as, the first and second moments of the lindley distribution in (5) are, respectively, as follows:

$$E(\lambda) = \frac{1}{1+\theta}(1+\frac{2}{\theta}) \text{ and } E(\lambda^2) = \frac{3+2\theta}{\theta^2(1+\theta)} \qquad (6)$$

2.3 NB-L Distribution

The effective formula in (1) for determining the number of times of ammunition faults in NB distributed processing is established, which has discrete count data. We combine the potential failure rate λ described by Lindley distribution with the parameter r of the NB distribution to obtain the NB-L mixture model, which is used to describe the number of ammunition malfunctions.

By integrating over all possible values of λ, we obtain the marginal probability of observing y failures. The PMF of the NB-L distribution is as follows

- $$P(Y=y;\theta)=\int_0^{\infty} NB(r,\mu)lindley(\theta)d\lambda = \frac{\Gamma(y+r)e^{-\theta\lambda}(1+\lambda)\theta^2}{\Gamma(y+1)\Gamma(r)(1+\theta)}\int_0^{\infty}\left(\frac{r}{\mu+r}\right)^{y_i}\left(\frac{\mu}{\mu+r}\right)^{r}d\lambda \tag{7}$$

The NB-L distribution directly reflects the distribution pattern of the final failure count caused by the rate heterogeneity depicted by the Lindley distribution. The distribution in Figure 2 can not only solve the problem of excessive zeros but also retain the fitting ability for non-zero observations.

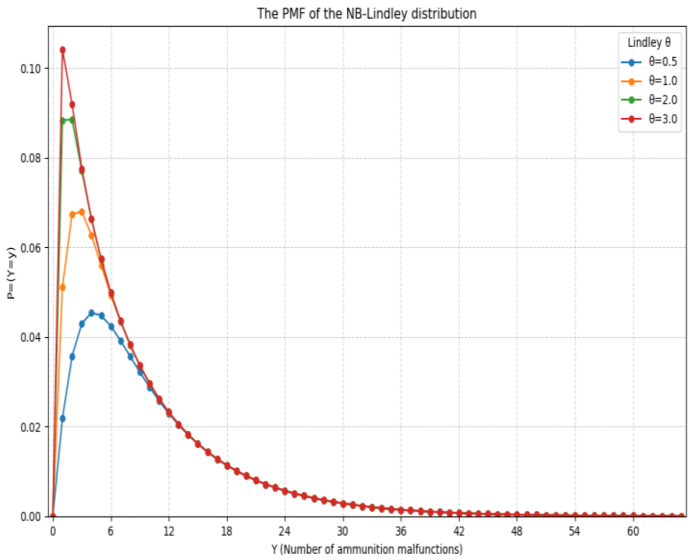

Figure 2 The PMF of the NB-Lindley distribution

Bayesian inference for the NB-L regression model

3.1 The GLM framework

Although the marginal distribution of failure λ counts follows the Lindley model, the characterization alone is insufficient for practical ammunition failure analysis. Quantifying the influence of explanatory variables on the mean failure intensity is essential. The generalized linear model consists of the following three components:

(1) Suppose that the response variable Y_i (the number of ammunition malfunctions for the i observation unit) follows the probability distribution related to the discrete parameter λ of the mean μ_i

(2) Define a linear predictor δ_i ,which is linear combination of covariates $X_i=(X_{i1},X_{i2},\cdots,X_{ik})$:

- $$g(\mu_i)=\ln(\mu_i)=\beta_0+\beta_1 x_{i1}+\cdots+\beta_p x_{ip} \tag{8}$$

where $X_{i1}, X_{i2}, \cdots, X_{ik}$ represents the factors influencing the risk of failure in the i-th observation and $\beta = (\beta_0, \beta_1, \cdots, \beta_p)^T$ represents the vector of regression coefficients to be estimated. β_0 represents intercept term, while $\beta_0, \beta_1, \cdots, \beta_p$ quantifies the influence of the p-th covariate on the risk of failure.

(3) Select a differentiable link function $g(\cdot)$,connect the average number of failures $\mu_i = E(Y_i)$ with the linear connector w_i,and establish a linear relationship between the response variable expectation Y_i The GLM framework of expression is:

- $$E(Y) = \mu_i = exp(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}) = exp\left(x_i^T \beta\right) \quad (9)$$

where μ_i represents the linear prediction factor, x_{ip} represents the vector of covariates, and β_j corresponds to the regression coefficient.

Suppose the conditional distribution $Y_i \mid x_i^T$ is $NB - L(\mu_i; r, \theta)$,

where $\mu_i = exp\left(x_i^T \beta\right)$.The PMF of $Y_i \mid x_i^T$ is obtained from Equations (1), (2) and (4) in the linear regression model ,that is:

- $$L(\Omega \mid y, x_i^T) = \frac{1}{(\Gamma(r))^n (1+\theta)^n} \prod_{i=1}^{n} \frac{\Gamma(y_i + r)}{\Gamma(y_i + 1)} \times \int_0^{\infty} \left(\frac{\lambda e^{x_i^T \beta}}{\lambda e^{x_i^T \beta} + r} \right)^{y_i} \left(\frac{r}{\lambda e^{x_i^T \beta} + r} \right)^{r} \left[\frac{(1+\lambda)\theta^2}{1+\theta} \right] e^{-\theta\lambda} d\lambda \quad (10)$$

where $y_i = 0, 1, 2, \ldots, \mu_i > 0, i = 1, 2, \ldots, n$, r, θ are positive parameters. Its mean and variance are respectively,

- $$E(Y_i \mid x_i^T) = \mu_i E(\lambda) = exp(\beta_0 + \sum_{i=1}^{\beta} X_i) \frac{\theta + 2}{\theta(\theta + 2)} \quad (11)$$
- $$Var(Y_i \mid x_i^T) = E(Y_i \mid x_i^T) + \mu_i^2 E(\lambda^2) \frac{1+r}{r} - E^2(Y_i \mid x_i^T) \quad (12)$$

Let $\Omega = (r, \theta, \beta)^T$ be the regression parameter vector, then the likelihood function of Ω is

- $$L(\Omega \mid y, x_i^T) = \frac{1}{(\Gamma(r))^n (1+\theta)^n} \prod_{i=1}^{n} \frac{\Gamma(y_i + r)}{\Gamma(y_i + 1)} \times \int_0^{\infty} \left(\frac{\lambda e^{x_i^T \beta}}{\lambda e^{x_i^T \beta} + r} \right)^{y_i} \left(\frac{r}{\lambda e^{x_i^T \beta} + r} \right)^{r} \left[\frac{(1+\lambda)\theta^2}{1+\theta} \right] e^{-\theta\lambda} d\lambda \quad (13)$$

3.2 Bayesian Hierarchical Model

The NB-L GLM model can be regarded as an extension form of GLM. Its likelihood function from Equation (17)) is not a closed-form expression but can be represented by a hierarchical Bayesian model.

Let the potential failure rate be $\lambda \sim Lindley(\theta)$,introducing an auxiliary latent variable ab, and it can be decomposed as

- $$\lambda \sim Lindley(\theta) \sim \frac{1}{1+\theta} gamma(2, \theta) + \frac{\theta}{1+\theta} gamma(1, \theta) \quad (14)$$

The hierarchical Bayesian framework of the NB-L GLM model adopts a three-level hierarchical structure. The mathematical expression is as follows:

$$f(y_i;\mu_i,r\mid\lambda)=NB(y_i;r,\lambda\mu_i)$$
$$\mu_i=exp\left(x_i^T\ \beta\right) \qquad (15)$$
$$\lambda\sim Lindley(\theta)$$

3.3 Prior Distributions And Joint Posterior Density

The Bayesian statistical framework systematically integrates prior information through setting probability distributions, and all unknown parameters are taken into account. The discrete parameters are usually assumed to follow the Gamma distribution in the negative binomial regression model（Gelman et al., 2013）, NB-L distribution proposed by Zamani and Ismail (2010) has the property that the parameter θ also follows gamma distribution. β represents the parameter of the GLM regression coefficient, while the normal distribution serves as the standard prior for the regression coefficient.

Let the parameters r,θ of the NB-L GLM model follow the gamma distribution, β follows the normal distribution, and all parameters are mutually independent. The joint prior distribution of the unknown parameters is as follows:

$$r\sim gamma(\alpha_r,z_r)$$
$$\theta\sim gamma(\alpha_\theta,z_\theta)$$
$$\beta\sim N(v_0,\sigma_\beta) \qquad (16)$$

where $\alpha_r,z_r,\alpha_\theta,z_\theta$ are a known positive parameter, v_0 is a vector of hyperparameters, and σ_β is a (k + 1)-order known non-negative specific matrix. The joint prior distribution of the unknown parameters is:

$$\pi(\Omega)=\pi(r)\pi(\theta)\pi(\beta) \qquad (17)$$

Owing to Bayes' theorem, the posterior distribution is obtained by multiplying the likelihood function by the prior distribution. Combining the likelihood function in Equation (14) and the prior distribution in Equation (18), the posterior distribution is obtained as:

$$\pi(\Omega\mid X)\propto L(\Omega\mid y,X)\pi(r)\pi(\theta)\pi(\beta) \qquad (18)$$

The full conditional posterior distributions for each parameter of X derived from (19) is obtained as

$$\pi(r\mid y,X,\theta,r)\propto L(\Omega\mid y,X)\pi(r)$$
$$\pi(\theta\mid y,X,\theta,r)\propto L(\Omega\mid y,X)\pi(\theta) \qquad (19)$$
$$\pi(\beta\mid y,X,\theta,r)\propto L(\Omega\mid y,X)\pi(\beta)$$

Empirical Application

4.1 Data Description

This section mainly selects the quality control project for ammunition production, aiming to analyze the factors influencing the number of ammunition malfunctions. The

data for this project were collected from multiple production facilities during the period from 2018 to 2022 for the batches of ammunition produced. Dataset contains 5190 observations, covering 5 different production batches. Each batch has 1000

-1010 observations, reflecting the diversity of different production scales and conditions.

Table 1 Summary statistics for the quantitative variables of Fault Occurrence Frequency

Variable	Max	Media	Min	Average(st)
Number	9	0	0	0.30
Gender	1	1	0	0.47
Age	0.72	0.32	0.19	0.40
Cost	1.5	0.55	0	0.58
Type	5	1	0	1.43
Measure	14	0.87	0	0.86
Production	12	0	0	1.22
Low-cost	1	0	0	0.44
Maintenanc	1	0	0	0.04

The number of faults that occurred was taken as the dependent variable, and several covariates were included, such as the production years of ammunition (measured in years), the gender of the operators (gender:1=female,0= male), the production cost divided by 10000(cost), the fault type(1=minor,5=severe), additional quality control measures (1=yes and 0=no), the production type (outsourcing: 1 = yes and 0=no), low-cost production (1=yes and 0=no), and free maintenance service (1=yes and 0=no). The descriptive summary of the variables presented in Table 1 indicates the zero proportion of the response variable number of failures is 0.798, the expected value of the number of failures is 0.30, the variance is 0.63, and the dispersion index is 2.11.

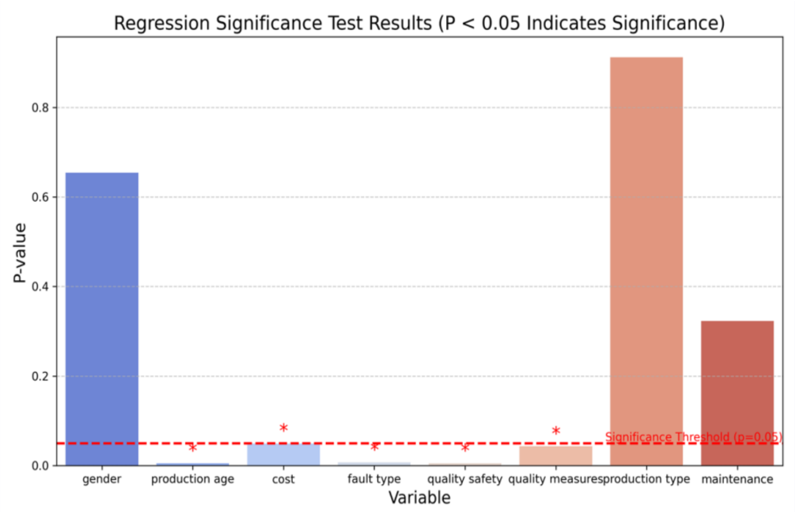

Figure 3 Results of Significance Test for Regression Analysis of Independent Variables

The preliminary analysis indicates that there is an excessive number of zero values for the number of faults. The summary of these data is presented in Table 1. Based on the statistical inference of the multiple linear regression model, combined with the VIF test and the analysis of variance method.

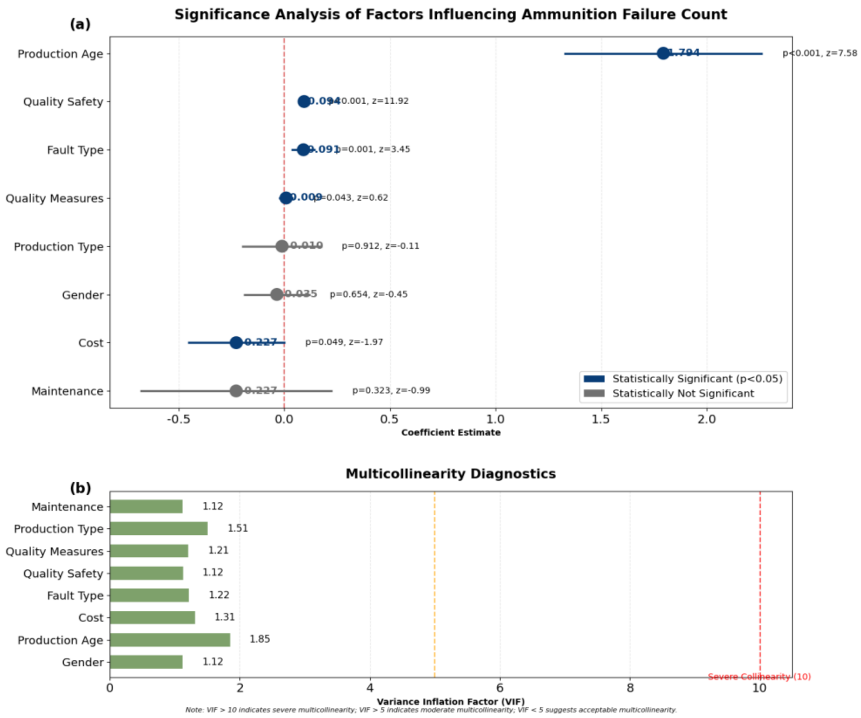

Figure 4 (a) Significance Analysis of Factors Affecting Ammunition Malfunctions (b) VIF Test for Multicollinearity Detection

As shown in Figures 3 and 4, for the multiple ammunition failure times affected by the forest plot presented and the significance test analysis, the p-values of the five variables "production age", "cost", "fault type", "quality safety" and " quality measure" are all less than 0.05, indicating that they are significant variables. Moreover, the larger the z-value, the greater influence of the factor on the number of ammunition failures. As shown in

Table 2 Summary statistics for the quantitative variables of Fault Occurrence

	Possion		NB		NB-L	
Parameter	Mean(s.e.)	95%Cr.I.	Mean(s.e.)	95%Cr.I.	Mean(s.e.)	95%Cr.I.
λ	0.30(0.05)	（0.68,0.89）	—	—	—	—
p	—	—	0.32(0.05)	(0.24,0.42)	—	—
r	—	—	0.37(0.07)	（0.26,0.52）	5.28(1.43)	(2.28,8.48)
θ	—	—	—	—	9.55(2.03)	(6.00,13.96)
Deviance	802.60		614.66		484.84	
DIC	803.61		616.52		490.50	
KS	0.0585		0.0135		0.0129	
p-value	0.0005		0.3009		0.3465	

Figure 4(b), it is no significant multicollinearity problem among the explanatory variables.
The variables that have been selected as significant ones for the quality control of ammunition production are "age","cost", "type", "control" and "measure".

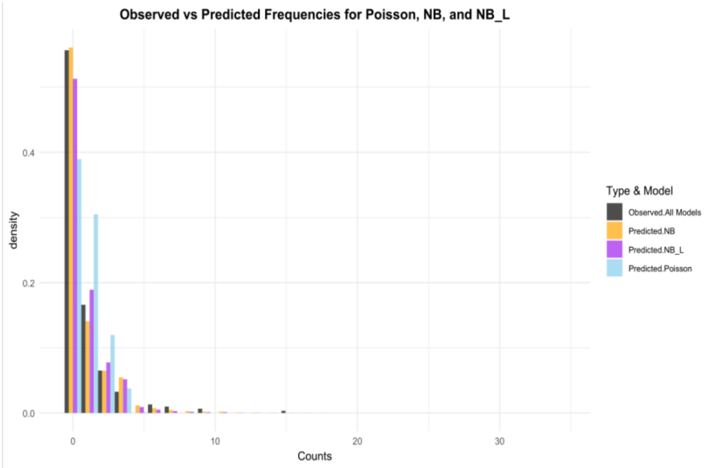

Figure 5 The bar chart plots of observed frequency and expected frequencies of each distributions.

4.2 Modeling Results

4.2.1 Bayesian Inference of NB-L Distribution

This part assesses the fitting performance of the NB-L distribution and compares it with the Poisson and the NB distribution. For each distribution, the Bayesian inference method is adopted to estimate the parameters. The KS test (Arnold and Emerson,2011), deviation and DIC were adopted to compare the fitting comparisons among the Poisson, NB and NB-L distributions. The best-fitting distribution for the data is the one that yields the minimum KS, deviation and DIC values.

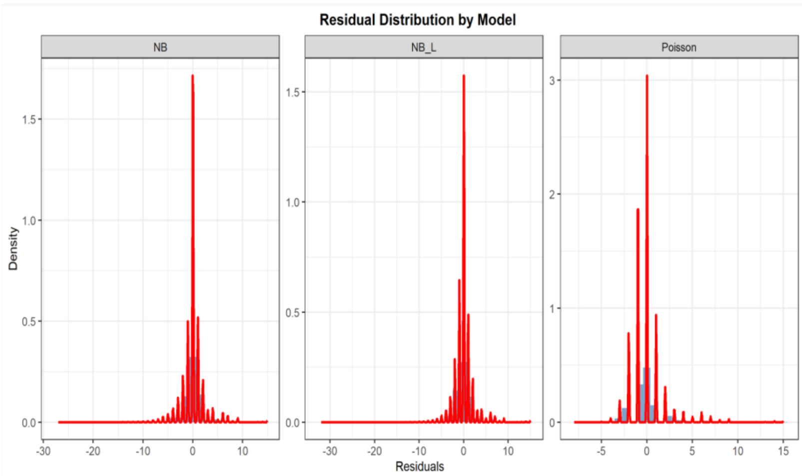

Figure 6 Visualization results of the residual distribution

Figure 5 presents the graphical comparison of the actual observed frequency distribution of the response variable "the number of failures" with theoretical frequency distributions predicted by the Poisson, NB and NB-L distributions.

The visualization results in Figure 6 indicate that compared with Poisson distribution, both NB and NB-L distribution can better fit the observed data. Among them, the goodness of fit of the NB-L distribution and the NB distribution to the actual observed values is particularly close.

Tables 2 present the posterior means, standard errors (s.e), 95% confidence intervals for each parameter, KS values, deviations and DIC values of the NB-L distribution, Poisson distribution and NB distribution.

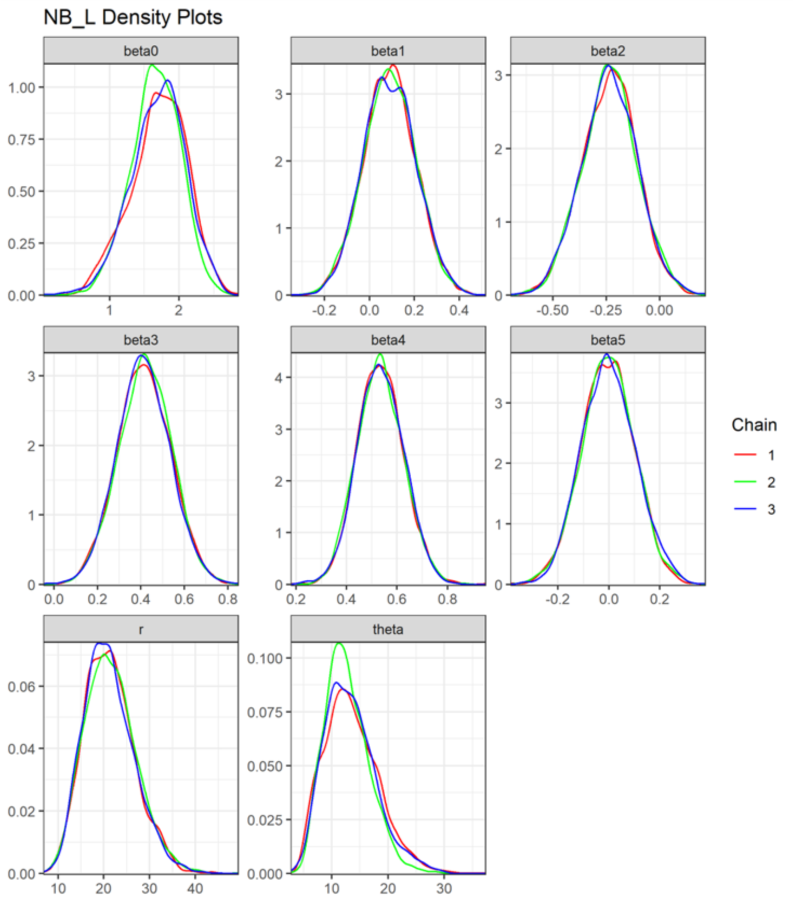

Figure 7 Density plots of the three MCMC chain for r, θ and $\beta = (\beta_0, \cdots \beta_5)^T$ from NB-L regression model

The results of the ammunition failure dataset in Table 2 show NB and NB-L distribution have better fitting performance than the Poisson distribution. Moreover, the KS value, deviation and DIC provided by the NB-L distribution are greater than those of the NB distribution.

However, these values of the NB-L distribution are close to the KS values of the NB distribution. Therefore, the NB-L distribution is a suitable distribution for describing this dataset, and its results are close to those of the NB distribution.

4.2.2 Bayesian Inference of the NB-L GLM Regression Model

Estimate parameters using Bayesian methods in Equation (19) and conduct model comparison for Poisson, NB and NB-L distribution. The evaluation indicators include bias, DIC and p_D. The p-value is close to the actual parameter quantity. The Deviation, DIC and the minimum value of p_D indicate that the model has the optimal fitting performance.

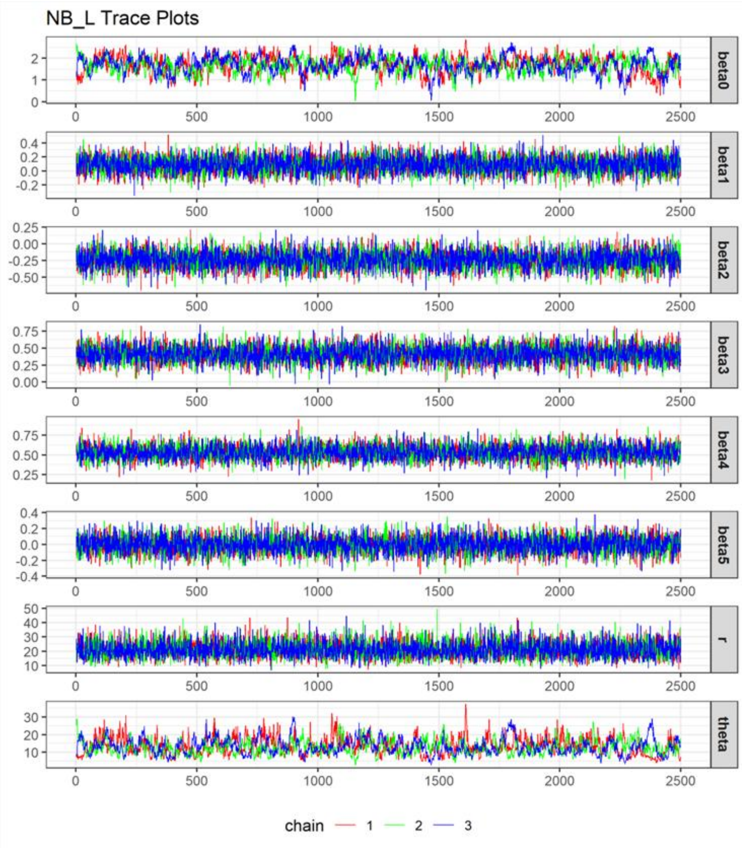

Figure 8 Trace plots of the three MCMC chain for r,θ and $\beta=(\beta_0,\cdots\beta_5)^T$ from NB-L regression model

According to Formula (4), incorporating the mean value of the GL distribution into the NB-L GLM

distribution for adjustment can enhance the fitting effect.

The Bayesian approach can be adopted to monitor and report the convergence of algorithms, which is a key aspect for ensuring the reliability of results. The sampling distribution of parameter values is analyzed by using the tracking graph and the posterior density graph to test the stationarity and convergence of the MCMC chain (Ntzoufras 2011).

Figure (7) graphically reflect the stability of the model during the sampling process, thereby ensuring the accuracy of the inference results.

Table 3 presents the parameter results of NB-L GLM, Position and NB regression models.

	Possion		NB		NB-L GLM	
Parameter	Mean(s.e.)	95%Cr.I.	Mean(s.e.)	95%Cr.I.	Mean(s.e.)	95%Cr.I.
β_0	-0.67(0.09)	(-0.85,-0.49)	-0.73(0.11)	(-0.95,0.52)	1.69(0.39)	(0.88,2.40)
β_1	0.01(0.08)	(-0.14,0.15)	0.08(0.11)	(-0.13,0.30)	0.09(0.12)	(-0.14,0.32)
β_2	0.28 (0.10)	(-0.47,-0.09)	-0.24(0.12)	(-0.48,0.00)	-0.24(0.13)	(-0.49,0.02)
β_3	0.31(0.08)	(0.16,0.46)	-0.39(0.11)	(0.18,0.62)	-0.42(0.12)	(0.17,0.65)
β_4	0.46(0.02)	(0.38,0.55)	0.52(0.08)	(0.37,0.69)	0.53(0.09)	(0.17,0.26)
β_5	0.04(0.06)	(-0.07,0.15)	-0.01(0.10)	(-0.20,0.18)	-0.01(0.10)	(-0.20,0.20)
r	—	—	1.50(0.58)	(0.76,2.96)	0.92(0.09)	(0.36,0.72)
θ	—	—	—	—	13.18(4.45)	(6.09,23.42)
Deviance	553.54		525.72		426.40	
DIC	559.66		532.67		432.78	

Among the data for analyzing ammunition malfunctions, a corresponding predictive model was formulated - it is related to five covariates of this data. The performance of the model is shown in Table 3.

The results show that the NB-L GLM model has the smallest deviance and DIC value, and the estimated value p_D is close to the implicit parameter β in the model.

The diagnostic plots for evaluating the regression model were considered, including posterior density plots and trace plots, to verify the applicability of the NB-L GLM model which can be generated in R language using the MCMCPLOTS. It is used to check the quality of the posterior distribution samples and the convergence of the model. The posterior density plot can intuitively display the characteristics of the parameter distribution, while the trace plot reflects the stability and consistency during the sampling process.

Figure (8) shows the posterior density plots of all parameters in the NB-LGLM model of the dataset. The results show that the posterior densities of the three parallel chains achieved a high degree of overlap after the burn period.

Concurrently, Figure () illustrates the changing trends of all parameters in the sampling sequence. The distribution of the simulated parameter values is dense and nearly vertical, which further validates the convergence and sampling stability of the model.

In conclusion, the NB-L regression model can fit this dataset from the results.

Conclusion

This paper focuses on the problem that there are excessive zero values and excessive dispersion in the data of ammunition failure times under combat conditions. The aim is to evaluate and verify the effectiveness and superiority of the NB-L distribution model in dealing with such complex count data. The research first expounded the theoretical basis of NB-L distribution and then constructed a Bayesian NB-L regression model based on the GLM framework. The introduction of Bayesian methods not only takes advantage of their strengths in handling small sample data and integrating prior information, but is applied through MCMC methods to estimate the parameters of NB-L distribution and the parameters of NB-L regression models. The reliability of the inference results was ensured through convergence diagnostics (trace plots and posterior density plots). Based on the deviation, the DIC and p-values results indicate that the NB-L GLM model has the best fitting effect, followed by the NB and Poisson regression models.

It can serve as an alternative model to study the relationship between the count response variable with excessive analysis and covariates. We hope that the GLM of NB-L distribution can replace the traditional Poisson

References

1. Altun, E. 2019. A new model for over-dispersed count data: Poisson quasi-Lindley regression model. Mathematical Sciences 13 (3):241–7. doi: 10.1007/s40096-019-0293.
2. Arnold, T., and J. Emerson. 2011. Nonparametric goodness-of-fit tests for discrete null distributions. The R Journal 3 (2):34–9. doi: 10.32614/RJ-2011-016.
3. Aryuyuen, S., and W. Bodhisuwan. 2013. The negative binomial-generalized exponential (NB-GE) distribution. Applied Mathematical Sciences 7(22):1093–105. doi:10.12988/ ams. 2013.13099.
4. Aryuyuen, S. 2021. Bayesian inference for the negative binomial-generalized Lindley regression model: properties and applications. Communications in Statistics - Theory and Methods 52(13):4534–4552. doi: 10.1080/036 10926.2021.1995434.
5. Croissant, Y., and S. Graves. 2020. Ecdat: Data sets for econometrics. R package version 0.3-9.https://CRAN.R-project.org/package= Ecdat.
6. 6. Dao, V. H., Gunawan, D., Kohn, R., Tran, M. N., Hawkins, G. E., Brown, S. D. 2025. Bayesian inference for evidence accumulation models with regressors. Psychological Methods. doi: 10.1037/met0000669.
7. Dzinyela, R., Shirazi, M., Das, S., and Lord, D. 2024. The negative Binomial-Lindley model with Time-Dependent Parameters: Accounting for temporal variations and excess zero observations in crash data. Accident Analysis & Prevention 207:107711. doi: 10. 1016/j.aap.2024.107711.
8. Feng, C. X. 2019. Zero-augmented accelerated

spatial failure model for modeling hospital length of stay data. Spatial and Spatio-temporal Epidemiology 29:121–137. doi: 10. 1016/j.sste.2018.05.001.

9. Fu, S. 2015. A hierarchical Bayesian approach to negative binomial regression. Methods and Applications of Analysis 22 (4):409–28. doi: 10.4310/MAA.2015.v22.n4.a4.
10. Fung, T. C. 2024. Robust estimation and diagnostic of generalized linear model for insurance losses: a weighted likelihood approach. Metrika 87 (3):333–366. doi: 10.10 07/s00184-023-00917-6.
11. Geedipally, S. R., Lord, D., and Dhavala, S. S. 2012. The negative binomial-Lindley generalized linear model: Characteristics and application using crash data. Accident Analysis and Prevention 45:258–65. doi: 10. 1016/j.aap.2011.07012
12. Ghitany, M. E., Atieh, B., and Nadarajah, S. 2008. Lindley distribution and its application. Mathematics and Computers in Simulation 78 (4):493–506. doi:10.1016/j.matcom.2007.06. 007.
13. Lord, D., Mannering, F. L. 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transportation Research – Part A 44 (5):291–305.
14. Lord, D., Geedipally, S. R. 2011. The negative binomial-Lindley distribution as a tool for analyzing crash data characterized by a large amount of zeros. Accident Analysis and Prevention 43 (5):1738–1742.
15. Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. 2013. The BUGS book: A practical introduction to Bayesian analysis. London: Chapman Hall.
16. Moksony, F., and Hegedűs, R. 2014. The use of Poisson regression in the sociological study of suicide. Corvinus Journal of Sociology and Social Policy 5 (2):97–114. doi: 10.14267/ cjssp.2014.02.04.
17. R Core Team. 2020. R: A Language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.
18. Su, S., and Yajima, M. 2021. R2jags: Using R to Run JAGS. R package version 0.7-1. https://CRAN.R-project.org/package=R2jags.
19. 19. Ünsal, Y. 2024. Economic complexity and environmental sustainability: sectoral perspectives from OECD countries. Journal of Environmental Studies and Sciences. doi: 10. 1007/s13412-024-00958-6.