OPEN ACCESS

Journal of Multidisciplinary Applied Natural Science Vol. xx No. xx (xxxx) https://doi.org/10.47352/jmans.2774-3047.444

Check for updates

Research Article

National Food Security Modeling Using the Zero-Inflated Ordered Probit with Correlated Errors (ZIOPC) Approach

Vita Ratnasari*, Nidya Putri Yudhani, I Nyoman Budiantara, and Ismaini Zain

Received : January 9, 2026 | Revised : March 31, 2026 | Accepted : April 14, 2026 | Online : June 22, 2026

Abstract

Zero-inflated ordinal data often emerge when the response variable displays both an ordered structure and an excessive number of zero outcomes. Such characteristics are not adequately captured by standard ordinal probit models. To address this limitation, this study employs a zero-inflated ordered probit with correlated errors (ZIOPC) model, which extends the conventional zero-inflated ordered probit (ZIOP) framework by explicitly modeling the dependence between the binary (inflation) and ordinal processes. From a modeling perspective, this study underscores the role of correlated latent errors in enhancing the representation of zero-inflated ordinal data. Parameter estimation was conducted using maximum likelihood estimation (MLE) implemented through the limited-memory BFGS with bound constraint (L-BFGS-B) algorithm to efficiently handle constrained optimization involving bivariate normal distributions. Statistical inference was performed using the maximum likelihood ratio test (MLRT) for simultaneous parameter testing, likelihood ratio test (LRT) for assessing the correlation parameter, and Wald test for partial significance of individual parameters. An empirical application using national food security data in Indonesia revealed substantial zero inflation. The results indicated that the correlation parameter was positive and statistically significant, confirming the presence of dependence between the two latent processes. Model comparison further demonstrates that the ZIOPC model provides a significantly better fit than the standard ZIOP model, as evidenced by the substantially lower AIC value of 1082.192 compared with 1187.468 for the ZIOP model. These findings emphasize the importance of incorporating correlated errors in modeling zero-inflated ordinal outcomes.

Keywords: food security, L-BFGS-B, ordinal Probit, ZIOPC

1. INTRODUCTION

Zero-inflated ordinal data frequently emerge in various applied disciplines, characterized by a response variable that exhibits both an ordered structure and an excessive number of zero observations [1]. Such data present significant challenges for conventional modeling approaches because standard ordinal regression models are not equipped to accommodate zero inflation [2]–[4]. To address this issue, the zero-inflated ordered probit (ZIOP) model was developed to capture the dual data-generating mechanism, comprising a binary component for structural zeros and an ordinal component for severity levels [5]–[8]. Despite its flexibility, the standard ZIOP model relies on the restrictive assumption that the error terms of the binary and ordinal components are independent. In many real-world scenarios, this assumption may be unrealistic, as unobserved factors can simultaneously influence both the probability of zero outcomes and the severity of the ordinal response [9][10]. Ignoring such dependence may result in biased parameter estimates and misleading inferences [11][12].

To overcome this limitation, this study employs the ZIOP with correlated errors (ZIOPC) model, which extends the ZIOP framework by allowing correlation between the latent error terms of the two components. This extension provides a more flexible and theoretically consistent approach for modeling zero-inflated ordinal data. Furthermore, this study emphasizes the numerical implementation of the ZIOPC model using maximum likelihood estimation (MLE) [13]–[15], combined with the limited-memory BFGS with bound constraints (L-BFGS-B) algorithm [16]. This approach is particularly suitable for addressing the constrained optimization problem arising from the presence of ordered thresholds and bounded correlation parameters, while avoiding the need for explicit second-order derivatives [17][18].

In addition to model estimation, this study also develops a comprehensive inference framework for the ZIOPC model. Specifically, parameter significance is evaluated using three complementary approaches. First, a maximum

Publisher's Note:
Pandawa Institute stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Please cite this article as: V. Ratnasari, N. P. Yudhani, I. N. Budiantara, and I. Zain. (2026). "National Food Security Modeling Using the Zero-Inflated Ordered Probit with Correlated Errors (ZIOPC) Approach". *Journal of Multidisciplinary Applied Natural Science*. https://doi.org/10.47352/jmans.2774-3047.444

Table 1. Variable used in the ZIOPC model.

Component	Variable	Description	Scale
Response variable	Z	National food security status 0: very secure 1: secure 2: somewhat secure 3: moderately vulnerable 4: vulnerable 5: severely vulnerable	Ordinal
Binary (inflation) component	X_1	Rice production	Ratio
	X_2	Percentage of poor population	Ratio
	X_3	Expenditure per capita	Ratio
Ordinal component	W_1	Gross regional domestic product	Ratio
	W_2	Harvested area	Ratio

likelihood ratio test (MLRT) is employed to assess the joint significance of model parameters [19]. Second, a likelihood ratio test (LRT) is used to examine the significance of the correlation parameter, thereby testing whether the ZIOPC model provides a statistically significant improvement over the standard ZIOP model. Third, Wald tests are conducted to evaluate the partial significance of individual parameters in both the binary and ordinal components. These inferential procedures ensure rigorous statistical validation of the proposed model and its components. As an empirical application, this framework was applied to the Indonesian national food security data. Food security remains a critical issue for social and economic stability [20], with regional disparities still evident despite recent improvements. The data exhibited a typical zero-inflated ordinal structure, making them an appropriate case for evaluating the performance of the ZIOPC model.

2. MATERIALS AND METHODS

2.1. Materials

This study employs secondary data sourced from the Badan Pusat Statistik (BPS) of Indonesia for 2024. The dataset encompasses 514 districts and municipalities and offers a comprehensive depiction of regional food security conditions. To investigate the determinants of food security status, three principal dimensions were considered: food availability, food access, and food utilization. Food availability is represented by rice production and harvested area, food access is proxied by gross regional domestic product and expenditure per capita, and food utilization is captured by the percentage of the poor population. These indicators were selected based on their pertinence to the conceptual framework of food security and availability in official statistics.

Within the ZIOPC modeling framework, the explanatory variables are categorized into two components: the binary (inflation) component and the ordinal component. This distinction reflects the different mechanisms underlying food security outcomes. The binary component models the probability of a region being structurally food secure (i.e., consistently in the “very secure” category), whereas the ordinal component captures the severity level of food insecurity among regions that are not structurally food secure.

Variables related to socioeconomic vulnerability, such as rice production, percentage of poor population, and expenditure per capita, are included in the binary components (X_1 - X_3) because they are anticipated to influence the likelihood of a region being structurally food secure. Conversely, variables reflecting regional economic capacity and agricultural structure, namely gross regional domestic product and harvested area, are included in the ordinal components (W_1 - W_2) because they are presumed to affect the degree of food insecurity severity. The definitions and measurements of all variables used in the ZIOPC model are presented in Table 1. This specification is grounded in theoretical considerations regarding the distinct roles of structural and severity mechanisms in food security analysis.

2.2. Methods

2.2.1. The ZIOPC Model

The ZIOPC model is designed to analyze ordinal response variables that exhibit excess zero outcomes. In this framework, the observed response may originate from two distinct mechanisms: a structural zero process and an ordinal severity process. The structural zero process determines whether an observation belongs to the inflated category, whereas the ordinal component determines the severity level when the observation arises from a non-inflated process. Let d be a binary indicator variable that determines whether an observation belongs to the ordinal component. The binary indicator is generated from a latent variable d^*, which is defined as Equation (1) [21];

$$d^* = \mathbf{x}^T\boldsymbol{\beta} + \varepsilon \quad (1)$$

Where $\mathbf{x}$ is the vector of explanatory variables associated with the binary component, $\boldsymbol{\beta}$ is the corresponding parameter vector, and ε is a random disturbance term. The observed binary variable is defined as Equation (2).

$$d = \begin{cases} 1, & if\ d^* > 0 \\ 0, & if\ d^* \leq 0 \end{cases} \quad (2)$$

Thus, the probability of belonging to the non-inflated component can be expressed as Equation (3);

$$P(d = 1 \mid \mathbf{x}) = \Phi\,(\mathbf{x}^T\boldsymbol{\beta}) \quad (3)$$

Where Φ(·) denotes the cumulative distribution function (CDF) of the standard normal distribution.

Conditional on $d = 1$, the ordinal outcome is generated through an ordered probit structure. Let s^* denote the latent variable associated with the ordinal response. The latent equation of the ordinal component is defined as Equation (4);

$$s^* = \mathbf{w}^T\boldsymbol{\gamma} + \delta \quad (4)$$

Where $\mathbf{w}$ is the vector of explanatory variables for the ordinal component, $\boldsymbol{\gamma}$ is the parameter vector non-inflated, and δ is the random error term.

The observed ordinal category s was obtained through a threshold mechanism applied to the latent variable s^*. Specifically, Equation (5) states

$$s = \begin{cases} 0, & s^* \leq \mu_0 \\ j, & \mu_{j-1} < s^* \leq \mu_j, \qquad j = 1, \dots, k-1 \\ k, & s^* > \mu_{k-1} \end{cases} \quad (5)$$

Where μ_0, μ_1, ..., μ_{k-1} are ordered threshold parameters satisfying $\mu_0 < \mu_1 < \dots < \mu_{k-1}$.

A key feature of the ZIOPC model is that the error terms of the binary and ordinal components can be correlated. Specifically, the disturbance terms are assumed to follow a bivariate normal distribution as listed in Equation (6);

$$(\varepsilon, \delta) \sim BVN\,(0, 0, 1, 1, \rho) \quad (6)$$

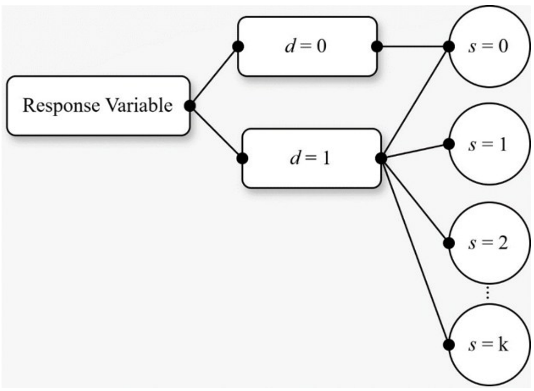

Figure 1. Structure of the ZIOPC model.

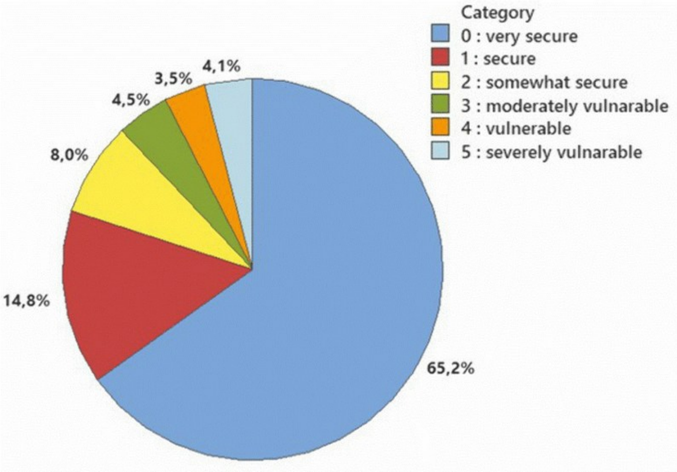

Figure 2. Percentage of national food security status in Indonesia.

Where ρ represents the correlation parameter capturing the dependence between the inflation mechanism and the ordinal severity process.

The observed response variable Z is defined as the product of the binary and ordinal components as Equation (7);

$$Z = d \times s \quad (7)$$

This formulation implies that zero outcomes may arise from two distinct sources, when $d = 0$ (structural zeros) or when $d = 1$ and $s = 0$ (sampling zeros from the ordinal process). Positive outcomes occur only when $d = 1$ and $s > 0$. This relationship is illusttrated in Figure 1.

Based on the model structure, the observed response variable can be expressed as Equation (8);

$$Z = \begin{cases} 0, & d = 0 \text{ or } (d = 1 \text{ and } s = 0) \\ j, & d = 1 \text{ and } s = j, \quad j = 1,2,\ldots,k-1 \\ k, & d = 1 \text{ and } s = k \end{cases} \quad (8)$$

Based on Equation (8), the conditional probability distribution of Z can be expressed as Equation (9);

$$P(Z = z|\mathbf{x},\mathbf{w}) = \begin{cases} P(Z = 0|\mathbf{x},\mathbf{w}) = P(d = 0|\mathbf{x}) + P(d = 1|\mathbf{x})P(s = 0|\mathbf{w}, d = 1; \rho) \\ P(Z = j|\mathbf{x},\mathbf{w}) = P(d = 1|\mathbf{x})P(s = j|\mathbf{w}, d = 1; \rho) \quad ; j = 1,2,\ldots,k-1 \\ P(Z = k|\mathbf{x},\mathbf{w}) = P(d = 1|\mathbf{x})P(s = k|\mathbf{w}, d = 1; \rho) \end{cases} \quad (9)$$

To derive closed-form expressions, the product terms in Equation (9) are equivalently expressed as joint probabilities. In particular,

$$P(d = 1 \,|\mathbf{x})P(s = j \,|\mathbf{w}, d = 1 \,; \rho) = P(d = 1, s = j \,|\mathbf{x},\mathbf{w})$$

which facilitates evaluation using the joint distribution of the error terms specified in Equation (6). From Equations (1) and (4), the event $d = 1$ corresponds to $\varepsilon > \mathbf{-x}^T\boldsymbol{\beta}$, while the ordinal condition $s = j$ corresponds to:

$$\mu_{j\text{-}1} < \mathbf{w}^T\boldsymbol{\gamma} + \delta \leq \mu_j$$

which can be written as follows.

$$\mu_{j\text{-}1} - \mathbf{w}^T\boldsymbol{\gamma} < \delta \leq \mu_j - \mathbf{w}^T\boldsymbol{\gamma}$$

Therefore, for $j = 1, 2, \ldots, k - 1$, the joint probability can be expressed as below.

$$P(d = 1, s = j \,|\mathbf{x}, \mathbf{w}) = P(\varepsilon > - \mathbf{x}^T\boldsymbol{\beta}, \mu_j - \mathbf{w}^T\boldsymbol{\gamma} < \delta \leq \mu_j - \mathbf{w}^T\boldsymbol{\gamma}$$

Using the properties of the bivariate normal distribution, this probability can be written as follows.

$$P(d = 1, s = j \,|\mathbf{x}, \mathbf{w}) = \Phi_2(\mathbf{x}^T\boldsymbol{\beta}, \mu_j - \mathbf{w}^T\boldsymbol{\gamma}; \rho) - \Phi_2(\mathbf{x}^T\boldsymbol{\beta}, \mu_{j\text{-}1} - \mathbf{w}^T\boldsymbol{\gamma}; \rho)$$

Similarly, for the lowest category,

$$P(d = 1, s = 0 \,|\mathbf{x}, \mathbf{w}) = P(\varepsilon > - \mathbf{x}^T\boldsymbol{\beta}, \; \delta \leq \mu_0 - \mathbf{w}^T\boldsymbol{\gamma}) = \Phi_2(\mathbf{x}^T\boldsymbol{\beta}, \mu_0 - \mathbf{w}^T\boldsymbol{\gamma}; \rho)$$

and for the highest category,

$$P(d = 1, s = k \,|\mathbf{x}, \mathbf{w}) = P(\varepsilon > - \mathbf{x}^T\boldsymbol{\beta}, \; \delta > \mu_{k\text{-}1} - \mathbf{w}^T\boldsymbol{\gamma}) = \Phi_2(\mathbf{x}^T\boldsymbol{\beta}, \mathbf{w}^T\boldsymbol{\gamma} - \mu_{k\text{-}1}; \rho)$$

Combining these results with $P(d = 0 \mid \mathbf{x}) = 1 - \Phi(\mathbf{x}^T\boldsymbol{\beta})$ the ZIOPC model can be expressed as:

$$P(Z = z|\mathbf{x},\mathbf{w}) = \begin{cases} P(Z = 0|\mathbf{x},\mathbf{w}) = \left(1 - \Phi(\mathbf{x}^T\boldsymbol{\beta})\right) + \Phi_2(\mathbf{x}^T\boldsymbol{\beta}, \mu_0 - \mathbf{w}^T\boldsymbol{\gamma}; \rho) \\ P(Z = j|\mathbf{x},\mathbf{w}) = \Phi_2(\mathbf{x}^T\boldsymbol{\beta}, \mu_j - \mathbf{w}^T\boldsymbol{\gamma}; \rho) - \Phi_2(\mathbf{x}^T\boldsymbol{\beta}, \mu_{j-1} - \mathbf{w}^T\boldsymbol{\gamma}; \rho); j = 1,2,\dots,k-1 \\ P(Z = k|\mathbf{x},\mathbf{w}) = \Phi_2(\mathbf{x}^T\boldsymbol{\beta}, \mathbf{w}^T\boldsymbol{\gamma} - \mu_{k-1}; \rho) \end{cases} \quad \textbf{(10)}$$

Where $\Phi(\cdot)$ is the CDF of the univariate standard normal distribution, $\Phi_2(\cdot)$ is the CDF of the bivariate normal distribution, and ρ represents the correlation between the error terms of the binary and ordinal components.

To ensure the identifiability of the ZIOPC model, several conditions must be satisfied. First, the scale of the latent variables is normalized by assuming that the error terms follow a standard normal distribution. Second, the threshold parameters are restricted to be strictly ordered, i.e., $\mu_0 < \mu_1 < \dots < \mu_{k-1}$, to preserve the ordinal structure of the response variable. Third, model identification is strengthened by allowing different sets of explanatory variables in the binary and ordinal components, which helps avoid collinearity between the two latent processes. Finally, the correlation parameter ρ is identifiable through the joint distribution of the error terms, as it captures the dependence between the inflation and ordinal mechanisms. These conditions ensure that the parameters of the ZIOPC model are uniquely determined and can be consistently estimated.

2.2.2. Estimating the ZIOPC Model Parameters

The parameters of the ZIOPC model are estimated using the MLE approach. Consider a random sample of size n, where the observed response variable Z_i takes values in the set $\{0, 1, \dots, k\}$. Conditional on the explanatory variables $\mathbf{x}_i$ and $\mathbf{w}_i$, the outcome Z_i follows a multinomial distribution given by

$$Z_i \sim M(1{:}P(Z_i = 0 \mid \mathbf{x}_i, \mathbf{w}_i), P(Z_i = 1 \mid \mathbf{x}_i, \mathbf{w}_i), \dots, P(Z_i = k \mid \mathbf{x}_i, \mathbf{w}_i)) \quad \textbf{(11)}$$

The joint probability distribution of the sample can be expressed as:

$$P(Z_1 = z_1, \dots, Z_n = z_n) = \prod_{i=1}^{n}\prod_{e=0}^{k} P(Z_i = e|\mathbf{x}_i, \mathbf{w}_i)^{h_{ei}} \quad \textbf{(12)}$$

Where h_{ei} is an indicator defined as

$$h_{ei} = \begin{cases} 1, & \text{if } Z_i = e \\ 0, & \text{otherwise} \end{cases} ; \; e = 0,1,\dots,k \quad \textbf{(13)}$$

Thus, the likelihood function of the ZIOPC model is given by:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n}\prod_{e=0}^{k} P(Z_i = e|\mathbf{x}_i, \mathbf{w}_i)^{h_{ei}} \quad \textbf{(14)}$$

where the parameter vector is defined as $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\mu}, \rho)$.

Taking the natural logarithm of Equation (14), the log-likelihood function is obtained as Equation (15).

$$\ln L(\boldsymbol{\theta}) = \sum_{i=1}^{n}\sum_{e=0}^{k} h_{ei} \ln P(Z_i = e|\mathbf{x}_i, \mathbf{w}_i) \quad \textbf{(15)}$$

The parameter estimates were obtained by maximizing the log-likelihood function in Equation (15). However, because of the presence of the bivariate normal CDF, $\Phi_2(\cdot, \cdot; \rho)$, the likelihood equations do not admit a closed-form solution. Therefore, numerical optimization methods were required. In this study, maximization was carried out using the L-BFGS-B algorithm [18]. This quasi-Newton method iteratively updates the parameter vector using gradient-based information while approximating the inverse Hessian matrix. The gradient of the log-likelihood function was computed numerically because of the involvement of the bivariate normal CDF. The L-BFGS-B

Table 2. Descriptive statistics of explanatory variables.

Variable	Mean	StDev	Minimum	Maximum
Rice production (X_1)	59.574	100.072	0.000	808.101
Percentage of poor population (X_2)	11.169	7.073	2.230	41.420
Expenditure per capita (X_3)	72.735	17.013	36.785	192.919
Gross regional domestic product (W_1)	262.571	560.994	1.735	5326.070
Harvested area (W_2)	19.545	30.413	0.000	212.866

Table 3. ZIOPC model parameter estimation results.

Predictor variable	*Coef*	*Std. error*	*P-value*
Binary (inflation) component			
Rice production (X_1)	-0.0051	0.0003	<0.0001*
Percentage of poor population (X_2)	-0.0386	0.0026	<0.0001*
Expenditure per capita (X_3)	0.0188	0.0128	0.1413
Constant	0.7737	0.8813	0.3800
Ordinal component			
Gross regional domestic product (W_1)	-0.0003	0.0002	0.0496*
Harvested area (W_2)	0.0116	0.0030	0.0001*
Correlation (ρ)	0.7810	0.0398	<0.0001*
Threshold (μ_0)	0.3213		
Threshold (μ_1)	0.7990		
Threshold (μ_2)	1.1730		
Threshold (μ_3)	1.5635		
Threshold (μ_4)	3.0957		

*Significant at $\alpha = 5\%$

procedure includes the following steps: (1) set the initial value $\boldsymbol{\theta}^{(0)} = \mathbf{0}$, a vector of dimension $(q + 1) + (p + 1) + (k - 1) + 1$, corresponding to the parameters $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\mu}$, and ρ, (2) set the initial value Hessian approximation $\mathbf{H}^{(0)} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, (3) compute the gradient vector $\mathbf{g}(\boldsymbol{\theta})^{(m)} = \nabla\ell(\boldsymbol{\theta})^{(m)}$, (4) compute the search direction as listed in Equation (16);

$$\mathbf{S}^{(m)} = [\mathbf{H}(\boldsymbol{\theta})^{(m)}]^{-1}\mathbf{g}(\boldsymbol{\theta})^{(m)} \tag{16}$$

The further steps are: (5) determine the step size $\alpha^{(m)}$ using line search:

$$\alpha^{(m)} = argmin[f(\boldsymbol{\theta}^{(m)} + \alpha^{(m)}\mathbf{S}^{(m)})] = \frac{[\mathbf{g}(\boldsymbol{\theta})^{(m)}]^T\mathbf{S}^{(m)}}{[\mathbf{S}^{(m)}]^{-1}\mathbf{H}(\boldsymbol{\theta})^{(m)}\mathbf{S}^{(m)}}$$

(6) Update the parameter vector:

$$(\boldsymbol{\theta})^{(m+1)} = (\boldsymbol{\theta})^{(m)} + \alpha^{(m)}\mathbf{S}^{(m)} \tag{17}$$

(7) Project the updated parameters onto the feasible set:

$$(\boldsymbol{\theta})^{(m+1)} = \Pi_c\left((\boldsymbol{\theta})^{(m+1)}\right) \tag{18}$$

Where $C = \{\boldsymbol{\theta} |\ -1 < \rho < 1,\ \mu_0 < \mu_1 < \dots < \mu_{k-1}\}$ with:

$$\rho^{(m+1)} = \min\left(\max(\rho^{(m+1)}, -1), 1\right)$$
$$\mu_j^{(m+1)} = \max(\mu_j^{(m+1)}, \mu_{j-1}^{(m+1)});\ j = 1, 2, \dots, k-1$$

(8) Update the inverse Hessian approximation using the BFGS formula:

$$\mathbf{H}(\boldsymbol{\theta})^{(m+1)} = \mathbf{H}(\boldsymbol{\theta})^{(m)} + \frac{\Delta(\boldsymbol{\theta})^{(m)}[\Delta(\boldsymbol{\theta})^{(m)}]^T}{[\Delta(\boldsymbol{\theta})^{(m)}]^T\Delta\mathbf{g}(\boldsymbol{\theta})^{(m)}} - \frac{\mathbf{H}(\boldsymbol{\theta})^{(m)}\Delta\mathbf{g}(\boldsymbol{\theta})^{(m)}[\Delta\mathbf{g}(\boldsymbol{\theta})^{(m)}]^T + \mathbf{H}(\boldsymbol{\theta})^{(m)}}{[\Delta\mathbf{g}(\boldsymbol{\theta})^{(m)}]^T\mathbf{H}(\boldsymbol{\theta})^{(m)}\Delta\mathbf{g}(\boldsymbol{\theta})^{(m)}} \tag{19}$$

Where $\Delta(\boldsymbol{\theta})^{(m)} = (\boldsymbol{\theta})^{(m+1)} - (\boldsymbol{\theta})^{(m)}$ and $\Delta\mathbf{g}(\boldsymbol{\theta})^{(m)} = \mathbf{g}(\boldsymbol{\theta})^{(m+1)} - \mathbf{g}(\boldsymbol{\theta})^{(m)}$.

(9) The iteration continues from step c until the convergence criterion is met:

$$\| (\boldsymbol{\theta})^{(m+1)} - (\boldsymbol{\theta})^{(m)} \| < \varepsilon \tag{20}$$

where the tolerance level is set to $\varepsilon = 10^{-5}$. All computations were implemented in Python using numerical optimization routines. The bivariate normal CDF was evaluated using stable numerical functions, and the L-BFGS-B algorithm was executed via standard scientific computing libraries.

2.2.3. Simultaneous Hypothesis Testing the ZIOPC Model Parameters

Simultaneous hypothesis testing was conducted using MLRT to evaluate the overall significance of the explanatory variables in the ZIOPC model. This test compares the log-likelihood of the full model with that of a restricted model under the null hypothesis. The hypotheses for testing the joint significance of the regression parameters are defined as

H_0: $\boldsymbol{\beta} = 0, \boldsymbol{\gamma} = 0$
H_1: at least one $\beta_u \neq 0$ or $\gamma_t \neq 0$; $u = 1, \ldots, q$; $t = 1, \ldots, p$ **(21)**

The MLRT statistic is given by:

$$LR = 2[\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_0)] \quad \textbf{(22)}$$

Where $\ell(\hat{\theta})$ is the maximized log-likelihood under the full model and $\ell(\hat{\theta}_0)$ is the maximized log-likelihood under the null hypothesis.

Under standard regularity conditions, the test statistic asymptotically follows a chi-squared distribution with degrees of freedom equal to the number of restrictions imposed under H_0, that is, $df = q + p$. The null hypothesis is rejected at significance level α if

$$LR > \chi^2_{\alpha,df} \quad \textbf{(23)}$$

To assess the significance of the correlation parameters, we conducted a separate LRT to compare the ZIOPC model with the standard ZIOP model. The hypotheses are defined as

H_0: $\rho = 0$
H_1: $\rho \neq 0$ **(24)**

The corresponding likelihood ratio statistic is given by

$$LR_\rho = 2[\ell(\hat{\theta}_{ZIOPC}) - \ell(\hat{\theta}_{ZIOP})] \quad \textbf{(25)}$$

Under the null hypothesis, LR_ρ asymptotically follows a chi-squared distribution with one degree of freedom. The null hypothesis is rejected if

$$LR_\rho > \chi^2_{\alpha,1} \quad \textbf{(26)}$$

2.2.4. Partial Hypothesis Testing the ZIOPC Model Parameters

Partial hypothesis testing was performed using the Wald test to evaluate the statistical significance of individual regression coefficients. For the parameter β_u, the hypotheses are defined as

H_0: $\beta_u = 0$
H_1: $\beta_u \neq 0$ **(27)**

The test statistic is given by

$$Z = \frac{\hat{\beta}_u}{\sqrt{\mathrm{Var}(\hat{\beta}_u)}} \quad \textbf{(28)}$$

Similarly, for the parameter γ_t, the hypotheses are:

$H_0 : \gamma_t = 0$
$H_1 : \gamma_t \neq 0$ **(29)**

with the corresponding test statistic:

$$Z = \frac{\hat{\gamma}_t}{\sqrt{\mathrm{Var}(\hat{\gamma}_t)}} \quad \textbf{(30)}$$

where the variance is obtained from the diagonal elements of the inverse Hessian matrix evaluated using MLE. The null hypothesis is rejected at significance level α if

$$|Z| > Z_{\alpha/2} \quad \textbf{(31)}$$

3. RESULTS AND DISCUSSIONS

3.1. Descriptive Statistics

Indonesia's national food security status for 2024 predominantly fell within the highest category on an ordinal scale comprising six levels: very secure (0), secure (1), somewhat secure (2), moderately vulnerable (3), vulnerable (4), and severely vulnerable (5). Figure 2 illustrates that 65.18% of districts and municipalities were classified as "very secure." This indicates significant zero inflation in the response variable. The remaining categories displayed varying proportions, suggesting regional disparities in food security.

Table 2 presents an overview of the characteristics and variability of the explanatory variables by presenting descriptive statistics, including the mean, standard deviation, minimum, and maximum values. These summary measures indicate significant heterogeneity across districts and municipalities concerning agricultural production, economic capacity, and poverty levels, which are anticipated to influence regional food security outcomes.

3.2. ZIOPC Model Estimation

The parameters of the ZIOPC model were estimated utilizing the MLE method, employing the L-BFGS-B algorithm. The results of this estimation are presented in Table 3.

Based on Equation (10), the estimated probability model can be expressed in terms of linear predictors. To simplify the notation, the linear predictors of the binary (inflation) and ordinal components are defined as $B = \mathbf{x}^T\hat{\boldsymbol{\beta}}$ and $O = \mathbf{w}^T\hat{\boldsymbol{\gamma}}$, respectively. Using the estimated parameters in Table 3, the linear predictors are given by

$$B = 0.7737 - 0.0051X_1 - 0.0386X_2 + 0.0188X_3$$
$$O = -0.0003W_1 + 0.0116W_2$$

The estimated threshold parameters are

$\mu_0 = 0.3213$, $\mu_1 = 0.7990$, $\mu_2 = 1.1730$, $\mu_3 = 1.5635$, $\mu_4 = 3.0957$

These components were substituted into the ZIOPC probability expressions to compute the category probabilities for each observation.

To illustrate the model, Surabaya City was considered with the following values:

$W_1 = 4854.49$, $W_2 = 1.32$, $X_1 = 4.24$, $X_2 = 3.96$, $X_3 = 106.14$

Substituting into the linear predictors yields:

$O = -0.0003(4854.49) + 0.0116(1.32) = -1.441$
$B = 0.7737 - 0.0051(4.24) - 0.0386(3.96) + 0.0188(106.14) = 2.596$

For example, the probability of $Z = 0$ is computed as:

$\mu_0 - O = 0.321 - (-1.441) = 1.762$
$P(Z = 0) = (1 - \Phi(2.596)) + \Phi_2(2.596; 1.762; 0.781) = 0.9646$

The probabilities for the other categories were computed similarly. These results indicate that Surabaya has a probability of 0.9646 of being classified in the very secure category, implying a high likelihood of structural food security.

Although this example illustrates the application of the ZIOPC model at the regional level, a comprehensive understanding of the determinants of food security requires statistical inference on the estimated parameters. Therefore, hypothesis testing was conducted to evaluate the overall significance of the model and the individual effects of the explanatory variables. Using the LRT, the obtained statistic is

$$LR = 26.295 > \chi^2_{0.05,5} = 11.07$$

Thus, the null hypothesis was rejected, indicating that the explanatory variables jointly have a significant effect on the model.

To assess the significance of the correlation parameter, an LRT comparing the ZIOPC and ZIOP models was conducted (see Table 4). The results indicate that the null hypothesis $H_0 : \rho = 0$ was rejected, confirming the presence of a significant dependence between the binary and ordinal components. To examine the contribution of individual predictors, partial hypothesis testing was conducted using the Wald test. The results of this test, as presented in Table 3, provide insights into the significance and direction of the effects of each explanatory variable within the binary and ordinal components.

3.2.1. Binary (Inflation) Component

The inflation component models the probability of belonging to the structural zero group, that is, regions that are always classified as very food

Table 4. Model comparison based on maximized log-likelihood, AIC, and LRT.

Model	Log-likelihood	AIC	LRT
ZIOP	-582.734	1187.468	107.276
ZIOPC	-529.096	1082.192	

secure. The results show that rice production and the percentage of the poor population are statistically significant, whereas expenditure per capita is not. The negative coefficient of rice production indicates that an increase in production reduces the probability of belonging to the structurally very secure group. This suggests that production alone does not ensure structural food security, as other factors, such as distribution inefficiencies and market instability, may still play a role [22][23]. Similarly, the percentage of the poor population has a negative and significant effect, implying that higher poverty levels decrease the likelihood of a region being structurally food secure. This highlights the importance of socioeconomic conditions in determining food security status [24].

3.2.2. Ordinal Component

The ordinal component captures the severity of food insecurity conditional on not belonging to the structural zero group. GRDP and harvested area were found to be statistically significant. The negative coefficient of GRDP indicates that higher economic output reduces the probability of transitioning to more severe food insecurity categories, suggesting improved resilience. In contrast, a positive coefficient of harvested area implies that larger agricultural land areas are associated with higher food insecurity severity. This indicates that agricultural expansion alone may not be sufficient to improve food security without improvements in productivity and resilience [23][25].

3.2.3. Correlation Parameter

The estimated correlation parameter ($\rho = 0.781$) is positive and statistically significant, indicating dependence between the binary and ordinal components. This implies that unobserved factors influencing structural food security are related to those affecting severity. This finding validates the use of the ZIOPC model, as it captures the dependence structure that cannot be represented under the standard ZIOP model [26][27].

3.2.4. Policy Implications

These results suggest that improving regional economic capacity and reducing poverty are essential for enhancing food security. Agricultural policies should emphasize not only land expansion, but also productivity and resilience improvements. The significant correlation between the structural and severity components further indicates that integrated policy interventions are required, because both dimensions are influenced by common underlying factors.

3.3. Model Comparison

To evaluate the relative performance of the competing models, a comparison between the ZIOP and ZIOPC models was conducted. The comparison was based on maximized log-likelihood values, LRT, and the Akaike information criterion (AIC). The comparison results are presented in Table 4.

The ZIOPC model achieves a substantially higher (less negative) log-likelihood value than the ZIOP model, indicating an improved fit to the observed data. This improvement arises from the inclusion of the correlation parameter, which allows for dependence between the binary (inflation) and ordinal components [21][26]. To formally assess whether this improvement is statistically significant, an LRT was conducted. The test statistic is computed as

$$LR = 2[-529.096 - (-582.734)] = 107.276$$

Under the null hypothesis $H_0 : \rho = 0$, the statistic follows a chi-square distribution with one degree of freedom. Since:

$$LR = 107.276 > \chi^2_{0.05,1} = 3.84$$

the null hypothesis was rejected. This result indicates that the correlation between the binary and ordinal components was statistically significant, thereby justifying the use of the ZIOPC model over the standard ZIOP model. In addition, the AIC value for the ZIOPC model (1082.192) is substantially lower than that for the ZIOP model (1187.468), further confirming its superior goodness-of-fit after accounting for model complexity. Overall, these results demonstrate that incorporating correlated errors provides a more accurate representation of the underlying data-generating process. The ZIOPC model improves the statistical fit and captures the dependence structure between food security incidence and severity, which cannot be accommodated by the ZIOP model. Therefore, the ZIOPC model is preferred for analyzing national food security conditions in Indonesia.

4. CONCLUSIONS

The national food security status in Indonesia in 2024 exhibits substantial zero inflation, with most districts classified in the highest category. This feature requires a modeling framework capable of jointly capturing structural zeros and ordinal severity. This study contributes by employing the ZIOPC model, which extends the standard ZIOP framework by explicitly allowing dependence between the binary (inflation) and ordinal components. In addition, this study highlights the practical implementation of the ZIOPC model using the L-BFGS-B optimization algorithm, which is particularly suitable for handling constrained and nonlinear likelihoods involving the bivariate normal distribution. The empirical results show that GRDP and harvested area significantly influence food insecurity severity, whereas rice production and the percentage of the poor population determine the probability of structural food security. The estimated correlation parameter is positive and statistically significant, confirming that the two latent processes are interdependent. This result demonstrates that relaxing the independence assumption leads to a more accurate representation of the data-generating process. Consistent with this finding, the model comparison results indicate that the ZIOPC model outperforms the ZIOP model. These findings emphasize the importance of jointly modeling structural and severity mechanisms in food security analysis. Future research may extend the ZIOPC framework by incorporating spatial or time-series structures and exploring alternative optimization strategies for high-dimensional zero-inflated ordinal models.

AUTHOR INFORMATION

Corresponding Author

Vita Ratnasari — Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya-60111 (Indonesia);

orcid.org/0000-0002-7579-4752

Email: vitaratnasari.its@gmail.com

Authors

Nidya Putri Yudhani — Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya-60111 (Indonesia);

orcid.org/0009-0008-2219-0315

I Nyoman Budiantara — Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya-60111 (Indonesia);

orcid.org/0000-0001-6572-4083

Ismaini Zain — Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya-60111 (Indonesia);

orcid.org/0000-0002-3716-3669

Author Contributions

Conceptualization, Methodology, Formal Analysis, V. R. and N. P. Y.; Software and Visualization, N. P. Y.; Supervision, V. R.; Validation, V. R., I. N. B., and I. Z.; Writing – Original Draft Preparation, N. P. Y.; Writing – Review & Editing, I. N. B., and I. Z.

Conflicts of Interest

The authors declare no conflict of interest.

ACKNOWLEDGEMENT

This research was supported by the Directorate of Research and Community Service, Directorate General of Research and Development, Ministry of Higher Education, Science, and Technology of the Republic of Indonesia with a grant number of 017/C3/DT.05.00/PL/2025 and 1325/PKS/ITS/2025.

DECLARATION OF GENERATIVE AI

Not applicable.

REFERENCES

[1] M. N. Harris and X. Zhao. (2007). "A Zero-Inflated Ordered Probit Model, with an Application to Modelling Tobacco Consumption". *Journal of Econometrics.* **141** (2): 1073-1099. 10.1016/j.jeconom.2007.01.002.

[2] D. Dale and A. Sirchenko. (2021). "Estimation of Nested and Zero-Inflated Ordered Probit Models". *The Stata Journal.* **21** (1): 3-38. 10.1177/1536867X211000002.

[3] D. Lambert. (1992). "Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing". *Technometrics.* **34** (1): 1-14. 10.2307/1269547.

[4] Y. Min and A. Agresti. (2005). "Random Effect Models for Repeated Measures of Zero-Inflated Count Data". *Statistical Modelling.* **5** (1): 1-19. 10.1191/1471082X05st084oa.

[5] A. Ghanbarnejad, A. Homayuni, Z. Hosseini, and A. Madani. (2022). "Smoking Behavior Among Students: Using Health Belief Model and Zero-Inflated Ordered Probit Model". *Tobacco and Health.* **1** (2): 74-82. 10.34172/thj.2022.12.

[6] H. Wang, Z. Liu, X. Wang, D. Huang, L. Cao, and J. Wang. (2022). "Analysis of the Injury-Severity Outcomes of Maritime Accidents Using a Zero-Inflated Ordered Probit Model". *Ocean Engineering.* **258** : 111796. 10.1016/j.oceaneng.2022.111796.

[7] J. Wu, W. Fan, and W. Wang. (2020). "A Zero-Inflated Ordered Probit Model to Analyze Hazmat Truck Drivers' Violation Behavior and Associated Risk Factors". *IEEE Access.* **8** : 110974-110985. 10.1109/ACCESS.2020.3001165.

[8] B. Neelon and D. Chung. (2017). "The LZIP: A Bayesian Latent Factor Model for Correlated Zero-Inflated Counts". *Biometrics.* **73** (1): 185-196. 10.1111/biom.12558.

[9] B. E. Bagozzi, D. W. Hill, W. H. Moore, and B. Mukherjee. (2015). "Modeling Two Types of Peace: The Zero-Inflated Ordered Probit (ZiOP) Model in Conflict Research". *Journal of Conflict Resolution.* **59** (4): 728-752. 10.1177/0022002713520530.

[10] C. Xu, S. Xu, C. Wang, and J. Li. (2019). "Investigating the Factors Affecting Secondary Crash Frequency Caused by One Primary Crash Using Zero-Inflated Ordered Probit Regression". *Physica A: Statistical Mechanics and Its Applications.* **524** : 121-129. 10.1016/j.physa.2019.03.036.

[11] R. Liesenfeld and J. F. Richard. (2010). "Efficient Estimation of Probit Models with Correlated Errors". *Journal of Econometrics.* **156** (2): 367-376. 10.1016/j.jeconom.2009.11.006.

[12] L. Cappellari and S. P. Jenkins. (2003). "Multivariate Probit Regression Using Simulated Maximum Likelihood". *The Stata Journal.* **3** (3): 278-294. 10.1177/1536867X0300300305.

[13] M. Zulfadhli, I. N. Budiantara, and V. Ratnasari. (2024). "Nonparametric Regression Estimator of Multivariable Fourier Series for Categorical Data". *MethodsX.* **13** : 102983. 10.1016/j.mex.2024.102983.

[14] M. Rifada, V. Ratnasari, and Purhadi. (2023). "Parameter Estimation and Hypothesis Testing of the Bivariate Polynomial Ordinal Logistic Regression Model". *Mathematics.* **11** (3): 579. 10.3390/math11030579.

[15] R. Guo, J. Zhu, X. Jiang, and F. Qu. (2025). "Fast and Efficient Implementation of the Maximum Likelihood Estimation for the Linear Regression with Gaussian Model Uncertainty". *Signal Processing.* 110403. 10.1016/j.sigpro.2025.110403.

[16] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu. (1995). "A Limited Memory Algorithm for Bound Constrained Optimization". *SIAM Journal on Scientific Computing.* **16** (5): 1190-1208. 10.1137/0916069.

[17] P. Venkataraman. (2009). "Applied Optimization with MATLAB Programming". John Wiley & Sons, New York.

[18] N. Andrei. (2018). "A Double Parameter Scaled BFGS Method for Unconstrained Optimization". *Journal of Computational and Applied Mathematics.* **332** : 26-44. 10.1016/j.cam.2017.10.009.

[19] V. Ratnasari, Purhadi, M. Rifada, and A. T. R. Dani. (2025). "Explore Poverty with Statistical Modeling: The Bivariate Polynomial Binary Logit Regression (BPBLR)". *MethodsX.* **14** : 103099. 10.1016/j.mex.2024.103099.

[20] H. Fransiska, A. M. Soleh, K. A. Notodiputro, and E. Erfiani. (2025). "Evaluation and Comparison of Three Mixed

-Effect Models for Household Food Insecurity Classification in West Java". *Journal of Multidisciplinary Applied Natural Science.* **6** (1): 466-480. 10.47352/jmans.2774-3047.334.

[21] W. H. Greene.(2020)." Econometric Analysis". Prentice Hall, New Jersey.

[22] C. B. Barrett. (2010). "Measuring Food Insecurity". *Science.* **327** (5967): 825-828. 10.1126/science.1182768.

[23] FAO, IFAD, UNICEF, WFP and WHO. (2023). "Urbanization, agrifood systems transformation and healthy diets across the rural–urban continuum". Rome, FAO. 10.4060/cc3017en

[24] C. Béné, D. Bakker, M. J. Chavarro, B. Even, J. Melo, and A. Sonneveld. (2021). "Global Assessment of the Impacts of COVID-19 on Food Security". *Global Food Security.* **31** : 100575. 10.1016/j.gfs.2021.100575.

[25] D. Laborde, W. Martin, J. Swinnen, and R. Vos. (2020). "COVID-19 Risks to Global Food Security". *Science.* **369** (6503): 500-502. 10.1126/science.abc4765.

[26] C. R. Bhat and S. K. Dubey. (2014). "A New Estimation Approach to Integrate Latent Psychological Constructs in Choice Modeling". *Transportation Research Part B: Methodological.* **67** : 68-85. 10.1016/j.trb.2014.04.011.

[27] Q. H. Vuong. (1989). "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses". *Econometrica.* **57** (2): 307-333. 10.2307/1912557.