

Article

Event-Driven Average Estimation with Dispersion-Tolerant Poisson–Inverse Gaussian Approach

Atef F. Hashem ¹, Asmaa S. Al-Moisheer ¹, Ahmet Bekir ², Ishfaq Ahmad ^{3,*} and Muhammad Raza ⁴¹ Department of Mathematics and Statistics, College of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia² Eskisehir Osmangazi University, Neighborhood of Akcaglan, Imarli Street, Number 28/4, 26030 Eskisehir, Turkey³ Department of Mathematics and Statistics, International Islamic University, Islamabad 44000, Pakistan⁴ PIDE School of Economics, Pakistan Institute of Development Economics (PIDE), Quaid-i-Azam University Campus, Islamabad 44000, Pakistan

* Correspondence: ishfaq.ahmad@iiu.edu.pk

Abstract

Overdispersion is a major problem in the context of count data analysis, and the classical Poisson regression estimators are, in general, unreliable since they imply that the mean equals its variance. In this article, an event-driven class of average estimators, which is based on the Poisson–inverse Gaussian (P-IG) regression model, is formulated to overcome this shortcoming. P-IG regression is a mixture of Poisson and inverse Gaussian regression that is modeled to deal with the overdispersion that is often found in real data. It approximates such count data by a compound distribution with a heavy-tailed inverse Gaussian component. Suggested estimators are more effective in estimating the population means in situations of overdispersion using auxiliary data in the form of covariates. The design-based framework specifies the statistical properties of proposed estimators with respect to their bias and mean squared error (MSE). To confirm the effectiveness and the strength of the suggested methodology, a reasonable amount of simulations and real-data applications are carried out, contrasting it with customary Poisson-based estimators. The results indicate that the P-IG-based estimators are superior over their counterparts. The study provides a statistically valid and practically useful breakthrough in survey sampling and count data regression that can provide researchers and practitioners with a strong alternative to classical Poisson-regression-based mean estimator procedures.

Keywords: event data; overdispersion; PIG regression; central tendency estimation**MSC:** 62D05

Academic Editors: Wenbo Wu and Chenlu Ke

Received: 27 September 2025

Revised: 18 November 2025

Accepted: 26 November 2025

Published: 28 November 2025

Citation: Hashem, A.F.; Al-Moisheer, A.S.; Bekir, A.; Ahmad, I.; Raza, M. Event-Driven Average Estimation with Dispersion-Tolerant Poisson–Inverse Gaussian Approach.

Mathematics **2025**, *13*, 3822. <https://doi.org/10.3390/math13233822>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

At every stage of a typical research process, conducting a sample survey is the initial step because it is a fundamental method for gathering data. Interviews may be conducted virtually or in person; arranged by individuals or groups; and conducted via mail, telephone, or in person methods. A questionnaire is one of the research instruments used across various disciplines, including education, medicine, economics, labor studies, industry studies, business studies, animal studies, and environmental studies. Personal, telephone, postal, and web-based surveys are the most common methods for conducting a survey. Surveys may be ineffective if flawed data collection methods are used. For

this reason, the methods of data collection should be accurate to yield valid conclusions. As a result, developing an action plan based on strategic data collection approaches can significantly increase the quality of research and enhance the accuracy of results. Utilizing auxiliary technical variables closely related to the primary research variable can enhance the model and yield more valid and stable results [1–3].

Sampling theory is distinct from other statistical theories in that it is based on supplementary information to maximize estimation accuracy. Auxiliary information is typically featured during different stages of a sampling study, including stratification and determination of selection probabilities, to maximize the effectiveness of population parameter estimation. Generally, the objective of surveys is to estimate the population total or mean. There are several techniques that have been suggested by researchers in using auxiliary information and improving the accuracy of population average estimation; see [4–7]. In the case of a direct relationship between supplementary and primary variables, regression and ratio estimators are particularly notable for estimating the population mean.

Nevertheless, classical techniques tend to become less accurate when extreme values or outliers in the data set come into play. To solve these shortcomings, effective ratio and regression-based techniques have been suggested in the estimation of means. Moreover, many sophisticated methods that specifically account for extreme values in augmented datasets have been designed. For example, Oral and Oral [8] treat the deficiencies of the classical ratio estimator and propose a new one constructed on the modified maximum likelihood (MML) and order statistics. Their estimator is very robust and efficient in non-normal cases. Abid et al. [9] explored estimating the population mean by employing unconventional location parameters. Zaman [10,11] and Bulut and Zaman [12] introduced a category of robust ratio estimators by synthesizing the existing ratio estimators. Ref. [13] and Ibrahim et al. [14] consider the problem of outliers in ratio estimators and employ robust regressions such as LTS, LMS, and Huber estimations. Under data contamination, their results indicate that estimation accuracy has been boosted by a substantial margin compared to the conventional approaches. In addition, the new double-sampling estimator, implied by robust regression methods, was proposed by Zaman and Bulut [15]. Ref. [16] suggested ratio-based estimators employing quantile regression as a replacement for conventional regression estimators. Ali et al. [17] obtained estimators through a robust regression technique, which is specifically appropriate in sensitive surveys that can be conducted in simple random samples. More recently, Koc [18] came up with the new idea of using the Poisson regression coefficient to estimate means in a specific situation, namely, in the context of count data. Their study is unique because they are concerned with count data and the utilization of Poisson regression in the estimation of ratio-based average estimators.

The basic principle of a useful statistical analysis has been count data modeling and estimation of mean. It can be directly applied in such domains as epidemiology, ecology, economics, and industrial engineering. The Poisson regression model of count outcomes has historically been the statistical workhorse because it is simple in both mathematics and analysis, and because it has a straightforward expression linking the mean of the response variable with its variance (Cameron and Trivedi [19]). That is one of the reasons that Koc [18] came up with the new idea of using the Poisson regression coefficient in mean estimation. After that, many studies extended their work by defining average estimators using the regression coefficient of Poisson. For example, Koc et al. [20] extend this work to double sampling. Ref. [21] developed a Poisson regression-based estimator of the mean that aims to suit the simple random sampling scenario to count data. Their approach is an indication of the significance of matching estimation methods with the distribution of underlying data. Wani et al. [22] proposed an enhanced version of the Koc [18] esti-

mator by incorporating power components from Shahzad et al. [21]. Raghav et al. [23] and Alomair and Shahzad [24] also extended this idea to Poisson and Tweedie–Poisson regression-based average estimators under probability sampling designed to improve the accuracy of count data.

However, the assumptions of the Poisson distribution do not always hold in the real world due to overdispersion in a data set. This violation may cause an underestimation of the standard errors and lead to inaccurate inference. To overcome these weaknesses, it is necessary to apply a more appropriate regression model that can accommodate dispersion. The P-IG regression model is another useful alternative, and it is a mixture of Poisson and inverse Gaussian. The model retains the simplicity of interpretation of the classical Poisson model and introduces a dispersion parameter to account for extra-Poisson variability; hence, it is particularly effective in fitting overdispersed counts (Sellers and Shmueli [25]; Farghali et al. [26]). The principal goal of the current research is to formulate mean estimators on the basis of the P-IG regression coefficient, which is more convenient than the classical Poisson estimators under conditions of overdispersion.

As far as we are aware, no prior research has investigated the problem of mean estimation in sampling theory using the P-IG regression coefficient. This gap is addressed in the present paper, where mean estimators based on the P-IG regression have been derived and tested, and their performance is compared with Poisson regression-based mean estimators. These estimators are based on auxiliary information, i.e., extra covariates or known quantities that enhance the efficiency and accuracy of statistical estimators (Sarndal et al. [27]; Rao [28]). The strategic application of auxiliary variables is a well-established procedure in survey sampling and small-area estimation; however, the incorporation of auxiliary variables into more sophisticated regression models involving count data has received relatively little attention. This aspect is also explored in the current research article.

The subsequent parts of the paper are arranged into four parts to discuss how the development and application of robust mean estimators concerning the P-IG regression coefficient are carried out. In Section 2, the background of count regression is reviewed, and the existing estimators developed by Koc [18] based on Poisson regression coefficients are explained. Section 3 contains the theory of the P-IG regression, its superiority over the classical equivalents, and the construction of the new family of average estimators, which are more efficient by utilizing auxiliary variable information. Section 4 presents a critical comparison of the developed and classical ones in both simulation and empirical settings using three real-world populations of mussel data and three artificial ones generated using the P-IG distribution. Section 4.3 summarizes the findings, and excellent results of the proposed estimators are indicated. Section 5 offers the conclusion of the study.

2. Poisson Regression for Count Data

It is necessary to understand the context of the main principles of count data regression to appreciate the usefulness of traditional models and the strategic importance of more adaptive ones. Count data are non-negative integer observations that appear in a wide variety of applied situations, such as counting the number of hospital visits, insurance claims, and ecological events that are recorded. The Poisson regression model has traditionally been the preferred choice for modelling such data because it is simple and has a clear interpretive framework. It presupposes that the response variable is Poisson distributed, the average of which is linked with explanatory variables in a log-linear fashion [19,29].

One of the most popular statistical methods of the count-type data analysis is the Poisson regression model. The method can be used when the dependent variable y_i for

$i = 0, 1, 2, \dots, n$ represents the frequency of events observed within a fixed duration or spatial domain. y_i considers Poisson distribution with

$$P(y_i; \lambda_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!}, \quad y_i = 0, 1, 2, \dots$$

where

$$\mathbb{E}(y_i) = \text{Var}(y_i) = \lambda_i.$$

The respective log-likelihood of the set of observations is developed in the following way:

$$\mathcal{L}(\lambda; \mathbf{y}) = \sum_{i=1}^n [y_i \log(\lambda_i) - \lambda_i - \log(y_i!)]$$

Suppose that $\mathbf{X}$ is a predictor matrix of dimension $n \times (p + 1)$, and $\mathbf{x}_i$ is the i th row of $\mathbf{X}$. The model relates λ_i with the predictors by use of the logarithmic link function:

$$\log(\lambda_i) = \gamma_i = \mathbf{x}_i^\top \mathbb{R}_p$$

where $\mathbb{R}_p = (\mathbb{R}_{p_0}, \mathbb{R}_{p_1}, \dots, \mathbb{R}_{p_p})^\top$ signifies regression coefficients.

The log-likelihood is differentiated with respect to the coefficient vector and equated to zero in order to estimate $\mathbb{R}_p$. The maximization of the likelihood leads to the following:

$$\sum_{i=1}^n (y_i - \exp(\mathbf{x}_i^\top \mathbb{R}_p)) \mathbf{x}_i = 0$$

Such a set of nonlinear equations is usually solved using an iterative numerical solution algorithm like the Newton–Raphson or the Fisher scoring algorithm [19,29].

Mean Estimators Using Poisson Regression

The use of linear regression models to analyze count data is generally inappropriate because negative values of predicted results can occur, and these models do not satisfy the assumptions of constant variance. Such issues can also adversely affect mean estimation. Koc [18] attempted to address this problem by proposing ratio estimators for mean estimation that incorporate Poisson regression coefficients.

The ratio-type methods of estimation are beneficial in a survey sampling context where a strong positive correlation exists between the supplementary variable and the research variable. The method was initially introduced by Cochran in the middle of the 1900s, and since then, it has developed to form an essential part of statistical inference. Although it was initially used extensively in agricultural research, it has since been found helpful in a variety of other scientific fields. Interested readers can consult [30–32] to know more about ratio estimators and their different variations. Koc's [18] class of ratio-type mean estimators is given below:

$$\bar{y}_{koc(i)} = \begin{cases} \{\bar{Y}_s + \hat{\mathbb{R}}_p(\bar{X}_p - \bar{X}_s)\} \bar{X}_p \bar{x}^{-1} \\ \{\bar{Y}_s + \hat{\mathbb{R}}_p(\bar{X}_p - \bar{X}_s)\} \bar{X}_C \bar{x}_C^{-1} \\ \{\bar{Y}_s + \hat{\mathbb{R}}_p(\bar{X}_p - \bar{X}_s)\} \bar{X}_R \bar{x}_R^{-1} \\ \{\bar{Y}_s + \hat{\mathbb{R}}_p(\bar{X}_p - \bar{X}_s)\} \bar{X}_{RC} \bar{x}_{RC}^{-1} \\ \{\bar{Y}_s + \hat{\mathbb{R}}_p(\bar{X}_p - \bar{X}_s)\} \bar{X}_{CR} \bar{x}_{CR}^{-1} \end{cases} \quad (1)$$

where

$$\left\{ \begin{array}{l} \bar{X}_C = \bar{X}_p + C_x, \quad \bar{x}_C = \bar{X}_s + C_x, \\ \bar{X}_{\mathbb{R}} = \bar{X}_p + \mathbb{R}_{2x}, \quad \bar{x}_{\mathbb{R}} = \bar{X}_s + \mathbb{R}_{2x}, \\ \bar{X}_{\mathbb{R}C} = \bar{X}_p \mathbb{R}_{2x} + C_x, \quad \bar{x}_{\mathbb{R}C} = \bar{X}_s \mathbb{R}_{2x} + C_x, \\ \bar{X}_{C\mathbb{R}} = \bar{X}_p C_x + \mathbb{R}_{2x}, \quad \bar{x}_{C\mathbb{R}} = \bar{X}_s C_x + \mathbb{R}_{2x} \end{array} \right. \quad (2)$$

$$MSE(\bar{y}_{koc(i)}) = \left\{ \begin{array}{l} \psi_{n_s, N_p} [S_y^2 + w_1^2 S_x^2 + 2\mathbb{R}_p w_1 S_x^2 + \mathbb{R}_p^2 S_x^2 - 2w_1 S_{xy} - 2\mathbb{R}_p S_{xy}] \\ \psi_{n_s, N_p} [S_y^2 + w_2^2 S_x^2 + 2\mathbb{R}_p w_2 S_x^2 + \mathbb{R}_p^2 S_x^2 - 2w_2 S_{xy} - 2\mathbb{R}_p S_{xy}] \\ \psi_{n_s, N_p} [S_y^2 + w_3^2 S_x^2 + 2\mathbb{R}_p w_3 S_x^2 + \mathbb{R}_p^2 S_x^2 - 2w_3 S_{xy} - 2\mathbb{R}_p S_{xy}] \\ \psi_{n_s, N_p} [S_y^2 + w_4^2 S_x^2 + 2\mathbb{R}_p w_4 S_x^2 + \mathbb{R}_p^2 S_x^2 - 2w_4 S_{xy} - 2\mathbb{R}_p S_{xy}] \\ \psi_{n_s, N_p} [S_y^2 + w_5^2 S_x^2 + 2\mathbb{R}_p w_5 S_x^2 + \mathbb{R}_p^2 S_x^2 - 2w_5 S_{xy} - 2\mathbb{R}_p S_{xy}] \end{array} \right. \quad (3)$$

where

$$w_i = \left\{ \begin{array}{l} (\bar{Y}_p)(\bar{X}_p)^{-1} \\ (\bar{Y}_p)(\bar{X}_p + C_x)^{-1} \\ (\bar{Y}_p)(\bar{X}_p + \mathbb{R}_{2x})^{-1} \\ (\bar{Y}_p \mathbb{R}_{2x})(\bar{X}_p \mathbb{R}_{2x} + C_x)^{-1} \\ (\bar{Y}_p C_x)(\bar{X}_p C_x + \mathbb{R}_{2x})^{-1} \end{array} \right. \quad (4)$$

The $\bar{y}_{koc(i)}$ class has an MSE in Equation (3) that depends on finite population correction (fpc) factor ψ_{n_s, N_p} containing population and sample sizes (N_p, n_s), population averages ($\bar{X}_p, \bar{Y}_p$), sample averages ($\bar{X}_s, \bar{Y}_s$), variances (S_y^2, S_x^2), CV (C_x), kurtosis ($\mathbb{R}_{2x}$), and co-variance (S_{xy}).

3. Formulation of P-IG Regression and Associated Estimators

3.1. P-IG Regression

Poisson regression is widely used; however, in some cases, real data may exhibit repeated failures under the condition that the variance is significantly greater than the mean, which is known as overdispersion. In such situations, the Poisson model, which assumes equality of the mean and variance, performs poorly; the parameters are estimated with bias; the standard errors are underestimated; and, consequently, inference tends to be unreliable (Hilbe [33]). Although adjustments such as quasi-Poisson and negative binomial models partially account for overdispersion, they often do so at the expense of interpretability. They are also limited in modelling data with both continuous and discrete components. This challenge led us to consider the P-IG regression for mean estimation. It is a mixture of Poisson and inverse Gaussian distributions. The P-IG regression preserves the count data structure while allowing for overdispersion and mixed distributions, which

are essential feature for robust mean estimation in complex data sets. It has a probability mass function (PMF) defined as follows:

$$P(Y = y) = \frac{\mu^y e^{-\lambda\mu}}{y!} \sqrt{\frac{2\lambda}{\pi}} \int_0^\infty t^{y-\frac{1}{2}} \exp\left[-\lambda\left(t + \frac{\mu^2}{2t}\right)\right] dt, \quad y = 0, 1, 2, \dots$$

The foregoing is indeed the mixed-Poisson formulation of the P-IG distribution [34–36]. There are also closed-form representations based on the modified Bessel function of the second kind, $K_{y+\frac{1}{2}}(\lambda\mu)$, and recursive probability formulae that can be expressed in a computational form [37,38]. The P-IG family extends the Poisson model by introducing a dispersion parameter λ that inflates the variance $\text{Var}(Y) = \mu + \mu^3/\lambda$, making it suitable for overdispersed count data. When

$$Y_i \sim \text{P-IG}(\mu_i, \lambda),$$

the mean will be μ_i and variance

$$\text{Var}(Y_i) = \mu_i + \frac{\mu_i^3}{\lambda}.$$

where λ is parameter of dispersion. The mean μ_i is expressed as a log-linear model of the explanatory variables $\mathbf{x}_i$:

$$\mu_i = \exp(\mathbf{x}_i^\top \Omega),$$

Ω is the regression coefficients vector. The fact that the mean response is always positive, as ensured by the use of the log-link, is also consistent with the count data setting. Maximum likelihood estimation (MLE) is used to determine the Ω and λ parameters. The log-likelihood has the following form:

$$\ell(\Omega, \lambda) = \sum_{i=1}^n \log\left[P(Y_i = y_i \mid \mu_i = \exp(\mathbf{x}_i^\top \Omega), \lambda)\right]$$

Replacing the PMF, the log-likelihood gets to

$$\ell(\Omega, \lambda) = \sum_{i=1}^n \left[y_i \log(\mu_i) - \log(y_i!) - \lambda\mu_i + \frac{1}{2} \log\left(\frac{2\lambda}{\pi}\right) + \log\left(\int_0^\infty t^{y_i-\frac{1}{2}} \exp\left[-\lambda\left(t + \frac{\mu_i^2}{2t}\right)\right] dt \right) \right].$$

The optimization over Ω and λ is performed numerically through the application of various numerical optimization algorithms such as the Newton–Raphson or BFGS. The optimum value of λ is then estimated and the final estimates of Ω and λ are obtained. This estimation step enables the P-IG model to be adaptable and to accommodate a wide range of dispersion and information structures, which is beneficial in particular when the Poisson model assumptions are overly constraining.

3.2. Proposed PIG Regression-Based Estimators

Utilization of auxiliary information in statistical models is a well-known mechanism to improve both the accuracy and effectiveness of parameter estimation processes and is essential in the context of mean estimation in count data analysis. In P-IG regression-based mean estimation, supplementary variables are linked with the research variable, which can greatly reduce the MSE of the estimators. Such variables can be included in the regression model directly, or they can be used during post-estimation calibration processes. They enhance the accuracy of the predicted means $\hat{\mu}_i$, especially in overdispersed or structurally heterogeneous data, whereby the standard Poisson regression is usually inadequate. The P-IG model, $\text{Var}(Y_i) = \mu_i + \frac{\mu_i^3}{\lambda}$, offers a flexible of variance structure that

allows the maximum use of the informational worthiness of nuisance variables and could possibly fit with complex data structure. The theoretical and practical importance of the strategic use of supplementary information in the P-IG architecture lies in guaranteeing the robustness of mean estimation based on the foundations of model-assisted estimation and small-area inference (Sarndal et al. [27]; Rao [28]). We suggest the following mean estimators based on P-IG-regression with the help of auxiliary information and following Alomair and Shahzad [24]:

$$\tilde{\mu}_j = \frac{n_s}{N_p} \bar{Y}_s + \left(1 - \frac{n_s}{N_p}\right) \cdot \frac{\bar{Y}_s + \Omega(\bar{X}_p - \bar{X}_s)}{\bar{h}_a \bar{X}_s + \bar{h}_b} \cdot (\bar{h}_a \bar{X}_p + \bar{h}_b) \cdot \exp\left(\frac{\theta_a(\bar{X}_p - \bar{X}_s)}{\theta_a(\bar{X}_p + \bar{X}_s) + 2\theta_b}\right) \quad (5)$$

Under the assumption of $\bar{Y}_s = \bar{Y}_p(1 + \zeta_0)$, $\bar{X}_s = \bar{X}_p(1 + \zeta_1)$ where $|\zeta_0|, |\zeta_1| \approx 0$, the expectations are

$$\left. \begin{aligned} E(\zeta_k) &= 0, \\ E(\zeta_0^2) &= \psi_{n_s, N_p} C_Y^2, \\ E(\zeta_1^2) &= \psi_{n_s, N_p} C_X^2, \\ E(\zeta_0 \zeta_1) &= \psi_{n_s, N_p} C_{YX}, \\ \psi_{n_s, N_p} &= \left(\frac{1}{n_s} - \frac{1}{N_p}\right) \end{aligned} \right\} \quad (6)$$

The simplification of Equation (5), along with its extension in terms of Equation (6), is obtained as

$$\tilde{\mu}_j = \frac{n_s}{N_p} \bar{Y}_p(1 + \zeta_0) + \left(1 - \frac{n_s}{N_p}\right) (\bar{Y}_p(1 + \zeta_0) - \Omega \bar{X}_p \zeta_1) (1 + \eta_a \zeta_1)^{-1} \cdot \exp(\eta_b^2 \zeta_1^2 - \eta_b \zeta_1) \quad (7)$$

Approximated to $O(n_s^{-1})$:

$$\begin{aligned} \tilde{\mu}_j &= \frac{n_s}{N_p} \bar{Y}_p(1 + \zeta_0) + \left(1 - \frac{n_s}{N_p}\right) (\bar{Y}_p(1 + \zeta_0) - \bar{Y}_p(\eta_a + \eta_b) + \Omega \bar{X}_p \zeta_1) \\ &\quad + (\bar{Y}_p(\eta_a^2 + \eta_a \eta_b + 1.5\eta_b^2) + \Omega \bar{X}_p(\eta_a + \eta_b)) \zeta_1^2 - \bar{Y}_p(\eta_a + \eta_b) \zeta_0 \zeta_1 \end{aligned} \quad (8)$$

The bias of $\tilde{\mu}_j$, after eliminating $\bar{Y}_p$ and taking expectation, is obtained up to order $O(n_s^{-1})$ as

$$\text{Bias}(\tilde{\mu}_j) = \psi_{n_s, N_p} \left(1 - \frac{n_s}{N_p}\right) \left[(\Omega \bar{X}_p(\eta_a + \eta_b) + \bar{Y}_p(\eta_a^2 + \eta_a \eta_b - 1.5\eta_b^2)) S_X^2 - \bar{Y}_p(\eta_a + \eta_b) C_{YX} \right] \quad (9)$$

where

$$\eta_a = \frac{\bar{h}_a \bar{X}_p}{\bar{h}_a \bar{X}_p + \bar{h}_b}, \quad \eta_b = \frac{\theta_a \bar{X}_p}{2(\theta_a \bar{X}_p + \theta_b)}$$

From Equation (8), after subtracting $\bar{Y}_p$, squaring and applying the expectation operator, the MSE of $\tilde{\mu}_j$ is obtained up to $O(n_s^{-1})$ as given below:

$$\text{MSE}(\tilde{\mu}_j) = \psi_{n_s, N_p} (S_Y^2 + \delta_j^2 S_X^2 - 2\delta_j S_{YX}) \quad (10)$$

where

$$\delta_j = \left(1 - \frac{n_s}{N_p}\right) (R(\eta_a + \eta_b) + \Omega)$$

Table 1 outlines the family members associated with the suggested class $\tilde{\mu}_j$. MSE of $\tilde{\mu}_j$ estimator in the set of $j = 1, 2, 3, 4$, and 5 , is

$$MSE(\tilde{\mu}_j) = \begin{cases} \psi_{n_s, N_p} (S_Y^2 + \delta_1^2 S_X^2 - 2\delta_1 S_{YX}) \\ \psi_{n_s, N_p} (S_Y^2 + \delta_2^2 S_X^2 - 2\delta_2 S_{YX}) \\ \psi_{n_s, N_p} (S_Y^2 + \delta_3^2 S_X^2 - 2\delta_3 S_{YX}) \\ \psi_{n_s, N_p} (S_Y^2 + \delta_4^2 S_X^2 - 2\delta_4 S_{YX}) \\ \psi_{n_s, N_p} (S_Y^2 + \delta_5^2 S_X^2 - 2\delta_5 S_{YX}) \end{cases} \quad (11)$$

Table 1. Variants within the proposed class $\tilde{\mu}_j$.

Estimators	Constant Values			
	$\hat{h}_a$	$\hat{h}_b$	θ_a	θ_b
$\tilde{\mu}_1 = \frac{n_s}{N_p} \bar{Y}_s + \left(1 - \frac{n_s}{N_p}\right) \frac{\bar{Y}_s + \Omega(\bar{X}_p - \bar{X}_s)}{\hat{h}_a \bar{X}_s + \hat{h}_b} (\hat{h}_a \bar{X}_p + \hat{h}_b) \exp\left(\frac{\theta_a(\bar{X}_p - \bar{X}_s)}{\theta_a(\bar{X}_p + \bar{X}_s) + 2\theta_b}\right)$	1	0	1	0
$\tilde{\mu}_2 = \frac{n_s}{N_p} \bar{Y}_s + \left(1 - \frac{n_s}{N_p}\right) \frac{\bar{Y}_s + \Omega(\bar{X}_p - \bar{X}_s)}{\hat{h}_a \bar{X}_s + \hat{h}_b} (\hat{h}_a \bar{X}_p + \hat{h}_b) \exp\left(\frac{\theta_a(\bar{X}_p - \bar{X}_s)}{\theta_a(\bar{X}_p + \bar{X}_s) + 2\theta_b}\right)$	1	C_X	1	0
$\tilde{\mu}_3 = \frac{n_s}{N_p} \bar{Y}_s + \left(1 - \frac{n_s}{N_p}\right) \frac{\bar{Y}_s + \Omega(\bar{X}_p - \bar{X}_s)}{\hat{h}_a \bar{X}_s + \hat{h}_b} (\hat{h}_a \bar{X}_p + \hat{h}_b) \exp\left(\frac{\theta_a(\bar{X}_p - \bar{X}_s)}{\theta_a(\bar{X}_p + \bar{X}_s) + 2\theta_b}\right)$	1	Ω_{1X}	1	0
$\tilde{\mu}_4 = \frac{n_s}{N_p} \bar{Y}_s + \left(1 - \frac{n_s}{N_p}\right) \frac{\bar{Y}_s + \Omega(\bar{X}_p - \bar{X}_s)}{\hat{h}_a \bar{X}_s + \hat{h}_b} (\hat{h}_a \bar{X}_p + \hat{h}_b) \exp\left(\frac{\theta_a(\bar{X}_p - \bar{X}_s)}{\theta_a(\bar{X}_p + \bar{X}_s) + 2\theta_b}\right)$	1	Ω_{2X}	1	0
$\tilde{\mu}_5 = \frac{n_s}{N_p} \bar{Y}_s + \left(1 - \frac{n_s}{N_p}\right) \frac{\bar{Y}_s + \Omega(\bar{X}_p - \bar{X}_s)}{\hat{h}_a \bar{X}_s + \hat{h}_b} (\hat{h}_a \bar{X}_p + \hat{h}_b) \exp\left(\frac{\theta_a(\bar{X}_p - \bar{X}_s)}{\theta_a(\bar{X}_p + \bar{X}_s) + 2\theta_b}\right)$	1	S_X	1	0

4. Numerical Illustration

This section carries out the comparative performance assessment of the average estimators based on the P-IG and Poisson regression coefficients using three empirical and three simulated datasets.

4.1. Application to Mussel Dataset

We employed the well-known mussel dataset that was initially used by Ref. [39] and later by Ref. [40] to briefly discuss the empirical performance of the suggested mean estimators derived through the P-IG regression coefficient. This dataset contains $N = 44$ observations and only one response variable, which is a count-based variable representing the number of species of mussels (y), along with several explanatory variables that depend on the geographical, ecological, and chemical parameters of river environments. As complementary variables among these explanatory variables, we used the river system indicators, i.e., the Alabama–Coosa (AC), Savannah (SV), and Apalachicola (AP) systems. Due to the count nature of the response, the P-IG regression coefficient is especially appropriate because it allows for the overdispersion that appears so often in ecological count data.

There are a number of reasons why this dataset as an empirical platform is suitable for our study: First, the assumptions used in the P-IG framework are satisfied by the discrete and overdispersed response (study) variable. Second, the data contains a variety of auxiliary covariates, which allow us to evaluate how extra information influences the accuracy of mean estimation.

To analyze the estimators, three populations was assumed from mussel data:

Population 1: Y = the number of mussel species and X = Alabama–Coosa river system (AC).

Population 2: Y = the number of mussel species and X = Savannah river system (SV).

Population 3: Y = the number of mussel species and X = Apalachicola river system (AP).

4.2. Simulation Design

In order to evaluate the behavior of $\tilde{\mu}_j$ in diverse scenarios, we systematically generated datasets according to the P-IG distribution that represents a flexible family of models adapted for overdispersed counts and possesses the properties of both a discrete and continuous part. The generated data contain 100 observations of (y, x) . The covariate x , which was supposed to approximate environmental or spatial gradients (e.g., distance to a source, pollution index), was produced by an integer-valued uniform random variable over the range $[1, 33]$ and ordered in descending order to reflect the monotonic decreasing trend observed in actual ecological field studies. The response variable y for three artificial populations were then modeled to depend inversely on x via a linear mean functions with added stochastic noises:

$$\mu_i = 20 - 0.4 x_i + \omega_i, \quad \omega_i \sim \mathcal{N}(0, 1), \quad i = 1, 2, \dots, 100,$$

$$\mu_i = 18 - 0.5 x_i + \omega_i, \quad \omega_i \sim \mathcal{N}(0, 1), \quad i = 1, 2, \dots, 100,$$

$$\mu_i = 22 - 0.49 x_i + \omega_i, \quad \omega_i \sim \mathcal{N}(0, 1), \quad i = 1, 2, \dots, 100,$$

In the above expression, μ_i indicates conditional mean of response variable y_i given covariate x_i , and the stochastic error component is given by ω_i . It was intended to make it positive, so μ_i values were truncated at 0.5 below. The P-IG distribution was then used to generate the outcome variable y_i of each population:

First simulated population (Population 4):

$$y_i \sim \text{P-IG}(\mu_i, \lambda = 5),$$

Second simulated population (Population 5):

$$y_i \sim \text{P-IG}(\mu_i, \lambda = 6),$$

Third simulated population (Population 6):

$$y_i \sim \text{P-IG}(\mu_i, \lambda = 4),$$

where $(\lambda = 5, 6, 4)$ are the dispersion parameters, corresponding to $\sigma = \sqrt{\frac{1}{\lambda}}$. It is an arrangement that reads count-like properties and overdispersion, which are common in environmental and biological data. The final values of the y_i were rounded to the nearest integer to approximate them as count data, which is typically obtained during field research. In addition to yielding a simulation framework with all the essential statistical properties of real-life ecological data, (P-IG)-based simulation frameworks enable the robust testing of mean estimation procedures in complex, non-Gaussian environments.

Based on reviewer's suggestion, we have also generated another population using negative binomial distribution (NBD). The response variable y for fourth artificial population was modeled to depend inversely on x as defined earlier via a linear mean function with added stochastic noise:

$$\mu_i = 20 - 0.4 x_i + \omega_i, \quad \omega_i \sim \mathcal{N}(0, 1), \quad i = 1, 2, \dots, 100.$$

The NBD was then used to generate the outcome variable y_i of artificial population:

Fourth simulated population (Population 7):

$$y_i \sim \text{NBD}(\mu_i, \lambda = 2).$$

Note that the results of real-life Populations 1–3, and artificial Populations 4–7 with replications 1000, are available in Tables 2–10.

4.3. Interpretation

As illustrated across Tables 2–10, the simulation-based findings and empirical-based findings of the proposed P-IG regression-based mean estimators, which are identified as $\tilde{\mu}_j$, show that compared to the traditional Poisson regression-based estimators, which are represented by $\bar{y}_{koc(i)}$, the former has a clear advantage in not only the overall MSE but also the PRE.

Table 2. MSE computation with mussel data.

Estimators	Population 1	Population 2	Population 3
$\bar{y}_{koc(1)}$	18.09802	17.93694	21.00994
$\bar{y}_{koc(2)}$	17.35626	16.44771	19.66512
$\bar{y}_{koc(3)}$	19.79844	18.33648	24.12391
$\bar{y}_{koc(4)}$	18.80512	52.82268	22.25344
$\bar{y}_{koc(5)}$	21.17745	18.56836	25.93548
$\tilde{\mu}_1$	8.801967	8.401536	9.253721
$\tilde{\mu}_2$	8.738982	8.289148	9.142953
$\tilde{\mu}_3$	8.780960	8.274641	9.215744
$\tilde{\mu}_4$	8.941959	8.430957	9.499182
$\tilde{\mu}_5$	8.207373	7.837018	8.467741

Table 3. PRE analysis with mussel data from Population 1.

n	$\bar{y}_{koc(1)}$	$\bar{y}_{koc(2)}$	$\bar{y}_{koc(3)}$	$\bar{y}_{koc(4)}$	$\bar{y}_{koc(5)}$
$\tilde{\mu}_1$	205.6134	197.1862	224.9320	213.6468	240.5991
$\tilde{\mu}_2$	207.0953	198.6074	226.5531	215.1866	242.3332
$\tilde{\mu}_3$	206.1053	197.6580	225.4701	214.1579	241.1747
$\tilde{\mu}_4$	202.3944	194.0991	221.4105	210.3020	236.8324
$\tilde{\mu}_5$	220.5094	211.4716	241.2275	229.1247	258.0296

Table 4. PRE analysis with mussel data from Population 2.

n	$\bar{y}_{koc(1)}$	$\bar{y}_{koc(2)}$	$\bar{y}_{koc(3)}$	$\bar{y}_{koc(4)}$	$\bar{y}_{koc(5)}$
$\tilde{\mu}_1$	213.4960	195.7703	218.2515	628.7264	221.0115
$\tilde{\mu}_2$	216.3906	198.4247	221.2107	637.2510	224.0080
$\tilde{\mu}_3$	216.7700	198.7725	221.5985	638.3683	224.4008
$\tilde{\mu}_4$	212.7510	195.0871	217.4899	626.5324	220.2402
$\tilde{\mu}_5$	228.8746	209.8721	233.9727	674.0151	236.9314

Table 5. PRE analysis with mussel data from Population 3.

n	$\bar{y}_{koc(1)}$	$\bar{y}_{koc(2)}$	$\bar{y}_{koc(3)}$	$\bar{y}_{koc(4)}$	$\bar{y}_{koc(5)}$
$\tilde{\mu}_1$	227.0432	212.5104	260.6942	240.4810	280.2708
$\tilde{\mu}_2$	229.7939	215.0850	263.8526	243.3945	283.6663
$\tilde{\mu}_3$	227.9788	213.3862	261.7685	241.4720	281.4257
$\tilde{\mu}_4$	221.1764	207.0191	253.9578	234.2669	273.0285
$\tilde{\mu}_5$	248.1175	232.2357	284.8920	262.8026	306.2856

Table 6. MSE using simulation study.

Estimators	Population 4	Population 5	Population 6	Population 7
$\bar{y}_{koc(1)}$	30.95435	19.09929	36.38091	30.25956
$\bar{y}_{koc(2)}$	30.49407	18.82019	35.86090	29.78202
$\bar{y}_{koc(3)}$	32.44886	19.99843	38.06561	31.80542
$\bar{y}_{koc(4)}$	31.31463	19.31701	36.78756	30.63287
$\bar{y}_{koc(5)}$	34.48946	21.21078	40.35775	33.90591
$\tilde{\mu}_1$	22.12762	13.95673	26.32689	21.04035
$\tilde{\mu}_2$	22.07544	13.91839	26.26496	20.98232
$\tilde{\mu}_3$	22.14098	13.96653	26.34273	21.05520
$\tilde{\mu}_4$	22.29310	14.07780	26.52301	21.22409
$\tilde{\mu}_5$	21.46685	13.46525	25.53984	20.30167

Table 7. PRE analysis with simulated data from Population 4.

n	$\bar{y}_{koc(1)}$	$\bar{y}_{koc(2)}$	$\bar{y}_{koc(3)}$	$\bar{y}_{koc(4)}$	$\bar{y}_{koc(5)}$
$\tilde{\mu}_1$	139.8901	137.8100	146.6442	141.5183	155.8661
$\tilde{\mu}_2$	140.2208	138.1358	146.9908	141.8528	156.2346
$\tilde{\mu}_3$	139.8057	137.7269	146.5557	141.4329	155.7721
$\tilde{\mu}_4$	138.8517	136.7871	145.5556	140.4678	154.7091
$\tilde{\mu}_5$	144.1960	142.0519	151.1580	145.8744	160.6638

Table 8. PRE analysis with simulated data from Population 5.

n	$\bar{y}_{koc(1)}$	$\bar{y}_{koc(2)}$	$\bar{y}_{koc(3)}$	$\bar{y}_{koc(4)}$	$\bar{y}_{koc(5)}$
$\tilde{\mu}_1$	136.8465	134.8467	143.2888	138.4065	151.9753
$\tilde{\mu}_2$	137.2234	135.2181	143.6835	138.7877	152.3939
$\tilde{\mu}_3$	136.7504	134.7521	143.1883	138.3093	151.8687
$\tilde{\mu}_4$	135.6695	133.6870	142.0565	137.2161	150.6683
$\tilde{\mu}_5$	141.8413	139.7686	148.5188	143.4582	157.5223

Table 9. PRE analysis with simulated data from Population 6.

n	$\bar{y}_{koc(1)}$	$\bar{y}_{koc(2)}$	$\bar{y}_{koc(3)}$	$\bar{y}_{koc(4)}$	$\bar{y}_{koc(5)}$
$\tilde{\mu}_1$	138.1892	136.2140	144.5883	139.7338	153.2948
$\tilde{\mu}_2$	138.5150	136.5351	144.9292	140.0632	153.6562
$\tilde{\mu}_3$	138.1061	136.1321	144.5014	139.6497	153.2026
$\tilde{\mu}_4$	137.1674	135.2068	143.5192	138.7005	152.1613
$\tilde{\mu}_5$	142.4477	140.4116	149.0440	144.0399	158.0188

Table 10. PRE analysis with simulated data from Population 7.

n	$\bar{y}_{koc(1)}$	$\bar{y}_{koc(2)}$	$\bar{y}_{koc(3)}$	$\bar{y}_{koc(4)}$	$\bar{y}_{koc(5)}$
$\tilde{\mu}_1$	143.8168	141.5471	151.1639	145.5910	161.1470
$\tilde{\mu}_2$	144.2146	141.9387	151.5820	145.9938	161.5928
$\tilde{\mu}_3$	143.7154	141.4473	151.0573	145.4884	161.0334
$\tilde{\mu}_4$	142.5718	140.3218	149.8553	144.3307	159.7520
$\tilde{\mu}_5$	149.0497	146.6974	156.6641	150.8885	167.0105

Table 2 displays the computed MSE values for three real-world populations based on the mussel dataset. The suggested estimators have much smaller MSEs compared to the traditional ones in all three populations. It is worth noting that in all the populations, $\tilde{\mu}_5$

has the minimal MSE, thus proving to be stronger under overdispersion and the presence of possible outliers. As an example in Population 3, where mussel species are estimated by taking AP as auxiliary variable, $\tilde{\mu}_5$ leads to an MSE of 8.467741 as opposed to 19.66512 with $\bar{y}_{koc(2)}$, emphasizing the practical benefit of using the P-IG-based approach. The PRE of each estimator based on the three populations considered is further expounded in Tables 3–5. In Population 1 (Table 3), $\tilde{\mu}_5$ attains the highest PRE (306.2856) which surpasses all of the conventional estimators. The same efficiency level can be seen in Populations 2 and 3, where $\tilde{\mu}_5$ was equal to PRE of 236.9314 and 258.0296, respectively (Tables 4 and 5). These results indicate the increased efficiency of estimation that can be obtained by employing the P-IG regression framework.

Additional evidence can be found in Table 6, which gives the results of the simulation where the MSE values were calculated using three synthetic populations. Once again, the proposed estimators are better, and $\tilde{\mu}_5$ again has the smallest MSE. In Population 6, the MSE of $\tilde{\mu}_5$ is 25.53984, and it is much lower than the 36.38091 as obtained by $\bar{y}_{koc(1)}$. This advantage is also proved by PRE metrics in Tables 7–10.

For summarizing all the results, visual comparisons are also provided in Figures 1–7. These findings, therefore, confirm the assertion that the suggested P-IG-based averaging estimators are not only conceptually sound but also empirically better. They offer an effective technique for calculating the mean, especially in environments that are overdispersed and contain misdirected values. The consistency of MSE and PRE measures favors the practical use of these estimators $\tilde{\mu}_j$ in survey sampling and count data analysis.

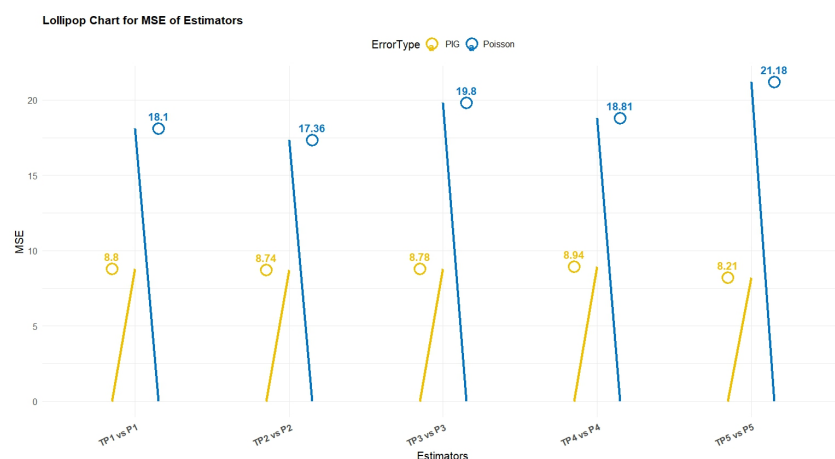


Figure 1. Performance evaluation based on Population 1.

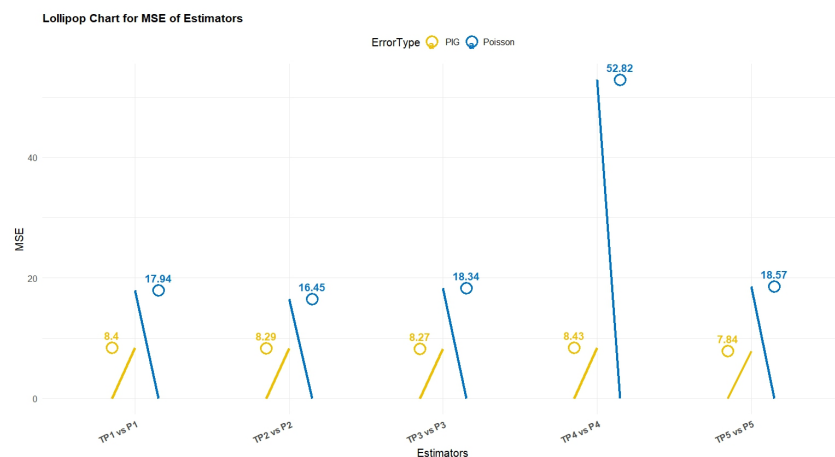


Figure 2. Performance evaluation based on Population 2.

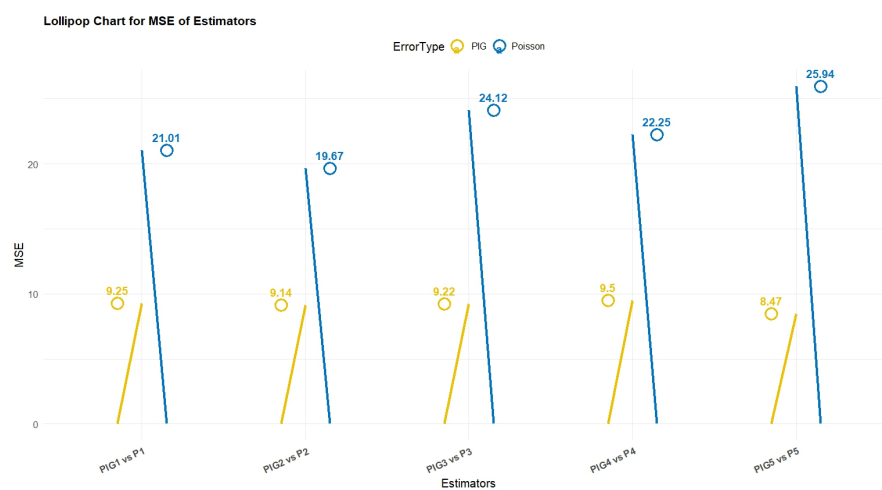


Figure 3. Performance evaluation based on Population 3.

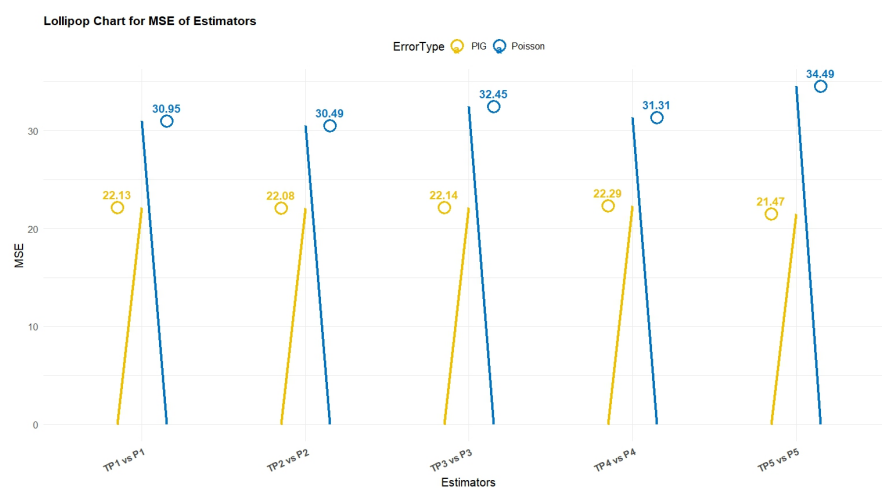


Figure 4. Performance evaluation based on Population 4.

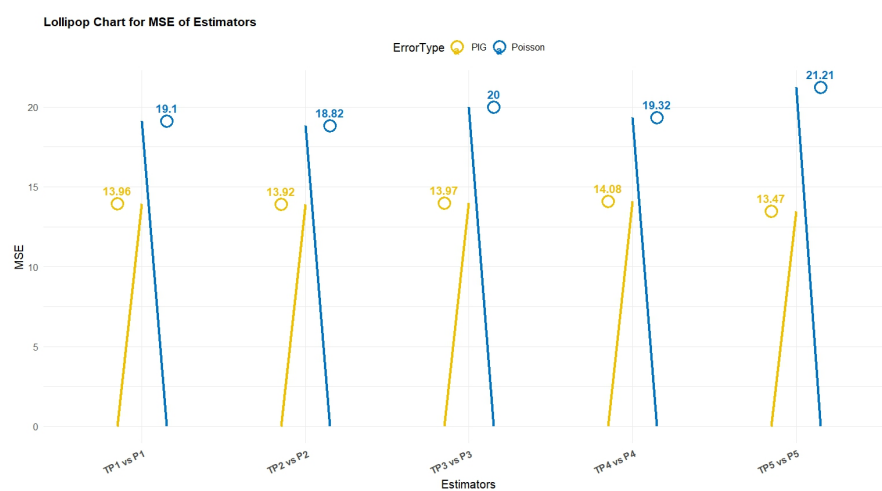


Figure 5. Performance evaluation based on Population 5.

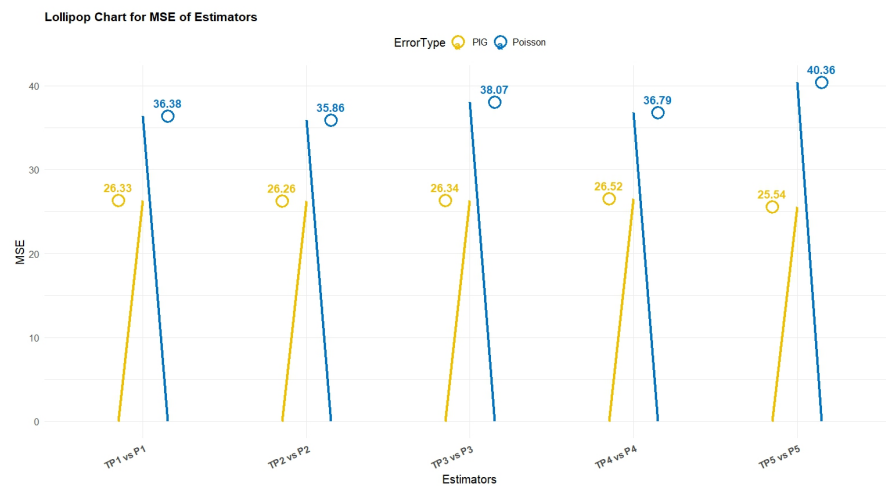


Figure 6. Performance evaluation based on Population 6.

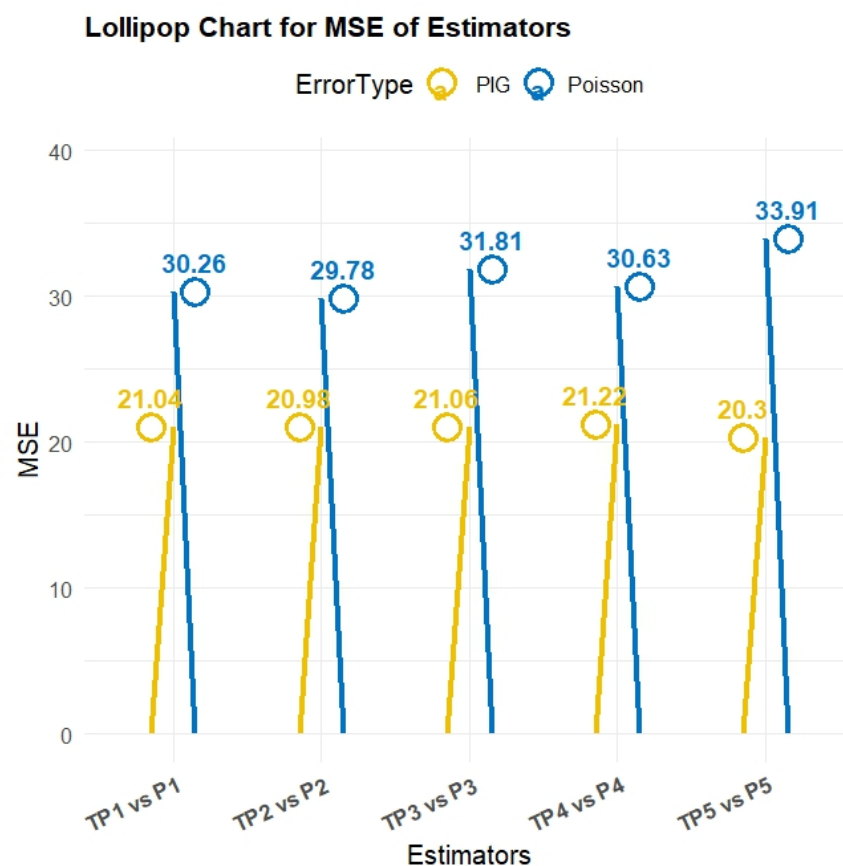


Figure 7. Performance evaluation based on Population 7.

5. Conclusions

The current article introduces an event-driven family of mean estimators that utilize the P-IG regression coefficient rather than relying on traditional Poisson regression-based estimators. The proposed family is empirically better and theoretically sound due to some crucial limitations of the Poisson model, which is incapable of dealing with overdispersed and heterogeneous count data. The theoretical derivations have revealed that the use of the supplementary variable that is implicit in the P-IG architecture can provide more efficient estimation in light of both MSE and PRE. Intensive tests on real-life data and synthesized data were used to support this theoretical contribution. To be precise, it was demonstrated in both the mussel dataset and in the simulated populations that the P-IG-based estimators

of the characteristic of the member of a population, namely the P-IG-based estimator $\tilde{\mu}_5$, were superior to the P-IG-based estimators by any of the utilized measures of estimator performance, i.e., smaller MSE and higher PRE value. The simulation studies also found these findings by applying the adaptability and robustness of the proposed estimators in complex data environments. These results reinforce the view that the mixed-type P-IG regression model offers a highly general and resilient framework for mean estimation under diverse dispersion and nonlinearity. Consequently, the integration of P-IG regression principles into the sampling theory represents a significant methodological contribution to statistical estimation.

Author Contributions: Conceptualization, A.F.H., A.B., I.A. and M.R.; methodology, A.B., I.A. and M.R.; software, A.F.H.; formal analysis, A.S.A.-M.; investigation, I.A.; resources, A.S.A.-M.; data curation, A.S.A.-M.; writing—original draft, A.F.H., A.S.A.-M., A.B., I.A. and M.R.; writing—review and editing, A.F.H., A.S.A.-M., A.B., I.A. and M.R.; project administration, I.A.; funding acquisition, A.F.H. and A.S.A.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP-RP25).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Diana, G.; Giordan, M.; Perri, P.F. An improved class of estimators for the population mean. *Stat. Methods Appl.* **2011**, *20*, 123–140. [\[CrossRef\]](#)
2. Diana, G.; Perri, P.F. A class of estimators in two-phase sampling with subsampling the non-respondents. *Appl. Math. Comput.* **2013**, *219*, 10033–10043. [\[CrossRef\]](#)
3. Ekpenyong, E.J.; Enang, E.I. Efficient exponential ratio estimator for estimating the population mean in simple random sampling. *Hacet. J. Math. Stat.* **2015**, *44*, 689–705. [\[CrossRef\]](#)
4. Singh, H.P.; Solanki, R.S. Improved estimation of population mean in simple random sampling using information on auxiliary attribute. *Appl. Math. Comput.* **2012**, *218*, 7798–7812. [\[CrossRef\]](#)
5. Riaz, S.; Darda, M.A. Some classes of estimators in the presence of non-response using auxiliary attribute. *SpringerPlus* **2016**, *5*, 1271. [\[CrossRef\]](#)
6. Subzar, M.; Maqbool, S.; Raja, T.A.; Pal, S.K.; Sharma, P. Efficient estimators of population mean using auxiliary information under simple random sampling. *Stat. Transit. New Ser.* **2018**, *19*, 219–238. [\[CrossRef\]](#)
7. Kumar, S.; Chhapparwal, P. Ratio-and product-based estimators using known coefficient of variation of the auxiliary variable via modified maximum likelihood. *Life Cycle Reliab. Saf. Eng.* **2019**, *8*, 99–116. [\[CrossRef\]](#)
8. Oral, E.; Oral, E. A robust alternative to the ratio estimator under non-normality. *Stat. Probab. Lett.* **2011**, *81*, 930–936. [\[CrossRef\]](#)
9. Abid, M.; Abbas, N.; Nazir, H.Z.; Lin, Z. Enhancing the mean ratio estimators for estimating population mean using non-conventional location parameters. *Rev. Colomb. Estad.* **2016**, *39*, 63–79. [\[CrossRef\]](#)
10. Zaman, T. Improvement of modified ratio estimators using robust regression methods. *Appl. Math. Comput.* **2019**, *348*, 627–631. [\[CrossRef\]](#)
11. Zaman, T. An efficient exponential estimator of the mean under stratified random sampling. *Math. Popul. Stud.* **2021**, *28*, 104–121. [\[CrossRef\]](#)
12. Bulut, H.; Zaman, T. An improved class of robust ratio estimators by using the minimum covariance determinant estimation. *Commun. Stat. Simul. Comput.* **2022**, *51*, 2457–2463. [\[CrossRef\]](#)
13. Subzar, M.; Bouza, C.N.; Al-Omari, A.I. Utilization of different robust regression techniques for estimation of finite population mean in SRSWOR in case of presence of outliers through ratio method of estimation. *Investig. Oper.* **2019**, *40*, 600–609.
14. Ibrahim, M.A.; Amer, I.A.; Emmanuel, J.E.M.S. Generalized Class of Mean Estimators with Known Measures for Outliers Treatment. *Comput. Syst. Sci. Eng.* **2021**, *38*, 1–15. [\[CrossRef\]](#)
15. Zaman, T.; Bulut, H. A simulation study: Robust ratio double sampling estimator of finite population mean in the presence of outliers. *Sci. Iran.* **2024**, *31*, 1330–1341. [\[CrossRef\]](#)

16. Shahzad, U.; Ahmad, I.; Al-Noor, N.H.; Hanif, M.; Almanjahie, I. Robust estimation of the population mean using quantile regression under systematic sampling. *Math. Popul. Stud.* **2023**, *30*, 195–207. [\[CrossRef\]](#)
17. Ali, N.; Ahmad, I.; Hanif, M.; Shahzad, U. Robust-regression-type estimators for improving mean estimation of sensitive variables by using auxiliary information. *Commun. Stat. Theory Methods* **2021**, *50*, 979–992. [\[CrossRef\]](#)
18. Koc, H. Ratio-type estimators for improving mean estimation using Poisson regression method. *Commun. Stat. Theory Methods* **2021**, *50*, 4685–4691. [\[CrossRef\]](#)
19. Cameron, A.C.; Trivedi, P.K. *Regression Analysis of Count Data*; Cambridge University Press: Cambridge, UK, 1998.
20. Koc, H.; Tanis, C.; Zaman, T. Poisson regression-ratio estimators of the population mean under double sampling, with application to COVID-19. *Math. Popul. Stud.* **2022**, *29*, 226–240. [\[CrossRef\]](#)
21. Shahzad, U.; Shahzadi, S.; Afshan, N.; Al-Noor, N.H.; Alilah, D.A.; Hanif, M.; Anas, M.M. Poisson Regression-Based Mean Estimator. *Math. Probl. Eng.* **2021**, *2021*, 9769029. [\[CrossRef\]](#)
22. Wani, Z.H.; Rizvi, S.E.H.; Sharma, M.; Jeelani, M.I.; Mushtaq, S. Modified Regression Estimators for Improving Mean Estimation-Poisson Regression Approach. *Pak. J. Stat. Oper. Res.* **2022**, *18*, 985–994.
23. Raghav, Y.S.; Ahmadini, A.A.H.; Mahnashi, A.M.; Rather, K.U.I. Enhancing estimation efficiency with proposed estimator: A comparative analysis of Poisson regression-based mean estimators. *Kuwait J. Sci.* **2025**, *52*, 100282. [\[CrossRef\]](#)
24. Alomair, A.M.; Shahzad, U. Tweedie-Poisson regression-based mean estimation of dose equivalence rates using gamma count auxiliary information. *J. Radiat. Res. Appl. Sci.* **2025**, *18*, 101929. [\[CrossRef\]](#)
25. Sellers, K.F.; Shmueli, G. A flexible regression model for count data. *Ann. Appl. Stat.* **2010**, *4*, 943–961. [\[CrossRef\]](#)
26. Farghali, R.A.; Lukman, A.F.; Algamal, Z.; Genc, M.; Attia, H. An Alternative Estimator for Poisson-Inverse-Gaussian Regression: The Modified Kibria-Lukman Estimator. *Algorithms* **2025**, *18*, 169. [\[CrossRef\]](#)
27. Sarndal, C.E.; Swensson, B.; Wretman, J. *Model Assisted Survey Sampling*; Springer: Berlin/Heidelberg, Germany, 1992.
28. Rao, J.N.K. *Small Area Estimation*; John Wiley and Sons: Hoboken, NJ, USA, 2003.
29. Saha, D.; Alluri, P.; Dumbaugh, E.; Gan, A. Application of the Poisson-Tweedie distribution in analyzing crash frequency data. *Accid. Anal. Prev.* **2020**, *137*, 105456. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Chutiman, N.; Nathomthong, A.; Wichitchan, S.; Guayjarernpanishk, P. Improved Estimator Using Auxiliary Information in Adaptive Cluster Sampling with Networks Selected Without Replacement. *Symmetry* **2025**, *17*, 375. [\[CrossRef\]](#)
31. Koc, T.; Koc, H. A new class of quantile regression ratio-type estimators for finite population mean in stratified random sampling. *Axioms* **2023**, *12*, 713. [\[CrossRef\]](#)
32. Wichitchan, S.; Nathomthong, A.; Guayjarernpanishk, P.; Chutiman, N. Enhanced Ratio-Type Estimators in Adaptive Cluster Sampling Using Jackknife Method. *Mathematics* **2025**, *13*, 2020. [\[CrossRef\]](#)
33. Hilbe, J.M. *Negative Binomial Regression*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2011.
34. Holla, M.S. On a Poisson-inverse Gaussian distribution. *Metrika* **1967**, *11*, 115–121. [\[CrossRef\]](#)
35. Sankaran, M. Mixtures by the inverse Gaussian distribution. *Sankhya B* **1968**, *30*, 455–458.
36. Ord, J.K.; Whitmore, G.A. The Poisson-inverse Gaussian distribution as a model for species abundance. *Commun. Stat. Theory Methods* **1986**, *15*, 853–871. [\[CrossRef\]](#)
37. Shaban, S.A. Computation of the Poisson-inverse Gaussian distribution. *Commun. Stat. Theory Methods* **1981**, *10*, 1389–1399. [\[CrossRef\]](#)
38. Ong, S.H. A note on the mixed Poisson formulation of the Poisson-inverse Gaussian distribution. *Commun. Stat. Simul. Comput.* **1998**, *27*, 67–78. [\[CrossRef\]](#)
39. Sepkoski, J.J., Jr.; Rex, M.A. Distribution of freshwater mussels: Coastal rivers as biogeographic islands. *Syst. Biol.* **1974**, *23*, 165–188. [\[CrossRef\]](#)
40. Batool, A.; Amin, M.; Elhassanein, A. On the performance of some new ridge parameter estimators in the Poisson-inverse Gaussian ridge regression. *Alex. Eng. J.* **2023**, *70*, 231–245. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.