

# Inference and Prediction in Inverse Gamma Regression: An Application to *Leptinaria unilamellata*

Tiago M. Magalhães<sup>1</sup>

(Joint work with Diego I. Gallardo<sup>2</sup>, Márcio A. Diniz<sup>3</sup> and Patrícia A. Daniel<sup>1</sup>)

<sup>1</sup>Federal University of Juiz de Fora, Brazil

<sup>2</sup>University of Bío-Bío, Chile

Icahn School of Medicine at Mount Sinai, USA<sup>3</sup>

Gramado, September 14, 2026



# Outline

- 1 Motivation
- 2 Inverse gamma regression model
- 3 Contributions
- 4 Simulation study
- 5 An application
- 6 Concluding remarks
- 7 References



# Outline

- 1 Motivation
- 2 Inverse gamma regression model
- 3 Contributions
- 4 Simulation study
- 5 An application
- 6 Concluding remarks
- 7 References



# Motivation

Regression analysis is perhaps the most widely used statistical technique (Montgomery et al., 2021). In many applications, the response variable is continuous, strictly positive, and right-skewed.

For instance, one may consider the time until offspring release in individuals of the species *Leptinaria unilamellata* (d'Orbigny, 1835), particularly under unfavorable environmental conditions where retention within the oviduct may occur.



# Motivation

Regression analysis is perhaps the most widely used statistical technique (Montgomery et al., 2021). In many applications, the response variable is continuous, strictly positive, and right-skewed.

For instance, one may consider the time until offspring release in individuals of the species *Leptinaria unilamellata* (d'Orbigny, 1835), particularly under unfavorable environmental conditions where retention within the oviduct may occur.



# Motivation

A common strategy in such settings is to transform the response variable in order to satisfy the assumptions of standard models, such as linear regression with normally distributed errors.

However, transformation-based approaches may be less appealing in applied contexts where direct interpretability is essential.



# Motivation

A common strategy in such settings is to transform the response variable in order to satisfy the assumptions of standard models, such as linear regression with normally distributed errors.

However, transformation-based approaches may be less appealing in applied contexts where direct interpretability is essential.



# Motivation

Defined on the positive real line, the inverse gamma distribution has several advantages (Magalhães et al., 2021) and its reparameterized counter part (RIG; Bourguignon and Gallardo, 2020) enables a regression framework in which both the mean and dispersion parameters are modelled directly.



# Motivation

The maximum likelihood (together with some of its variants) is the most widely used method of estimation (Lehmann and Casella, 1998).

However, the inferences derived from this method depend strongly on asymptotic properties of the maximum likelihood estimators.



# Motivation

The maximum likelihood (together with some of its variants) is the most widely used method of estimation (Lehmann and Casella, 1998).

However, the inferences derived from this method depend strongly on asymptotic properties of the maximum likelihood estimators.

Among these properties, we can highlight that these estimators follow a normal distribution, as the sample size increases.



# Motivation

The maximum likelihood (together with some of its variants) is the most widely used method of estimation (Lehmann and Casella, 1998).

However, the inferences derived from this method depend strongly on asymptotic properties of the maximum likelihood estimators.

Among these properties, we can highlight that these estimators follow a normal distribution, as the sample size increases.



# Motivation

The goal of this work is two-fold:

- to derive the skewness coefficient;



# Motivation

The goal of this work is two-fold:

- to derive the skewness coefficient;
- to obtain an effective procedure to compute prediction intervals.



# Motivation

The goal of this work is two-fold:

- to derive the skewness coefficient;
- to obtain an effective procedure to compute prediction intervals.

From Daniel (2013), we illustrate the proposed method by analyzing a data set in which the response variable is the intraoviductal retention duration in individuals of the species *Leptinaria unilamellata* (d'Orbigny, 1835).



# Motivation

The goal of this work is two-fold:

- to derive the skewness coefficient;
- to obtain an effective procedure to compute prediction intervals.

From Daniel (2013), we illustrate the proposed method by analyzing a data set in which the response variable is the intraoviductal retention duration in individuals of the species *Leptinaria unilamellata* (d'Orbigny, 1835).



# Outline

- 1 Motivation
- 2 Inverse gamma regression model
- 3 Contributions
- 4 Simulation study
- 5 An application
- 6 Concluding remarks
- 7 References



# Reparameterized inverse gamma distribution

We say that a random variable  $Y$  follows a RIG distribution with mean  $\mu > 0$  and precision parameter  $\phi > 0$ , if the distribution of  $Y$  admits the following probability density function (PDF):

$$f(y; \mu, \phi) = \frac{[\mu(1 + \phi)]^{2+\phi}}{\Gamma(2 + \phi)} y^{-(3+\phi)} \exp\left\{-\frac{\mu(1 + \phi)}{y}\right\} \mathbb{I}_{\mathbb{R}_+}(y), \quad (1)$$



# Inverse gamma distribution

where  $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$  is the complete gamma function, and  $\mathbb{I}_{\mathbb{R}_+}(y)$  is the indicator function that takes the value 1 if  $y \in \mathbb{R}_+$  is true and 0 otherwise.

The RIG differs from the inverse gamma distribution as the latter is a particular case of the proper dispersion models (Jørgensen, 1997), which is not mean-parameterized.



# Inverse gamma distribution

where  $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$  is the complete gamma function, and  $\mathbb{I}_{\mathbb{R}_+}(y)$  is the indicator function that takes the value 1 if  $y \in \mathbb{R}_+$  is true and 0 otherwise.

The RIG differs from the inverse gamma distribution as the latter is a particular case of the proper dispersion models (Jørgensen, 1997), which is not mean-parameterized.



# RIG regression model

Let  $Y_1, \dots, Y_n$  be  $n$  independent random variables, where each  $Y_i$ ,  $i = 1, \dots, n$ , follows the PDF given in (1) with mean  $\mu_i$  and precision parameter  $\phi_i$ .

Suppose the mean and precision parameters of  $Y_i$  satisfy the following functional relations:  $g_1(\mu_i) = \eta_{1i} = \mathbf{x}_i^\top \boldsymbol{\beta}$  and  $g_2(\phi_i) = \eta_{2i} = \mathbf{z}_i^\top \boldsymbol{\lambda}$ ,



# RIG regression model

Let  $Y_1, \dots, Y_n$  be  $n$  independent random variables, where each  $Y_i$ ,  $i = 1, \dots, n$ , follows the PDF given in (1) with mean  $\mu_i$  and precision parameter  $\phi_i$ .

Suppose the mean and precision parameters of  $Y_i$  satisfy the following functional relations:  $g_1(\mu_i) = \eta_{1i} = \mathbf{x}_i^\top \boldsymbol{\beta}$  and  $g_2(\phi_i) = \eta_{2i} = \mathbf{z}_i^\top \boldsymbol{\lambda}$ ,



# RIG regression model

where  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_{p_1})^\top$  and  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_{p_2})^\top$  are vectors of unknown regression coefficients which are assumed to be functionally independent,  $\boldsymbol{\beta} \in \mathbb{R}^{p_1}$  and  $\boldsymbol{\lambda} \in \mathbb{R}^{p_2}$ , with  $p_1 + p_2 < n$ ,  $\eta_{1i}$  and  $\eta_{2i}$  are the linear predictors, and  $\mathbf{x}_i = (x_{i1}, \dots, x_{ip_1})^\top$  and  $\mathbf{z}_i = (z_{i1}, \dots, z_{ip_2})^\top$  are observations on  $p_1$  and  $p_2$  known regressors, for  $i = 1, \dots, n$ .



# RIG regression model

Furthermore, we assume that the covariate matrices  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$  and  $\mathbf{Z} = (\mathbf{z}_1, \dots, \mathbf{z}_n)^\top$  have rank  $p_1$  and  $p_2$ , respectively.

The link functions  $g_1 : \mathbb{R} \rightarrow \mathbb{R}^+$  and  $g_2 : \mathbb{R} \rightarrow \mathbb{R}^+$  must be strictly monotone, positive, and at least twice differentiable such that  $\mu_i = g_1^{-1}(\mathbf{x}_i^\top \boldsymbol{\beta})$  and  $\phi_i = g_2^{-1}(\mathbf{z}_i^\top \boldsymbol{\lambda})$ , with  $g_1^{-1}(\cdot)$  and  $g_2^{-1}(\cdot)$  being the inverse functions of  $g_1(\cdot)$  and  $g_2(\cdot)$ , respectively.



# RIG regression model

Furthermore, we assume that the covariate matrices  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$  and  $\mathbf{Z} = (\mathbf{z}_1, \dots, \mathbf{z}_n)^\top$  have rank  $p_1$  and  $p_2$ , respectively.

The link functions  $g_1 : \mathbb{R} \rightarrow \mathbb{R}^+$  and  $g_2 : \mathbb{R} \rightarrow \mathbb{R}^+$  must be strictly monotone, positive, and at least twice differentiable such that  $\mu_i = g_1^{-1}(\mathbf{x}_i^\top \boldsymbol{\beta})$  and  $\phi_i = g_2^{-1}(\mathbf{z}_i^\top \boldsymbol{\lambda})$ , with  $g_1^{-1}(\cdot)$  and  $g_2^{-1}(\cdot)$  being the inverse functions of  $g_1(\cdot)$  and  $g_2(\cdot)$ , respectively.



# RIG regression model

For the log-likelihood function, the score vector and a detailed survey about maximum likelihood estimation, we suggest the original work of Bourguignon and Gallardo (2020).



# Outline

- 1 Motivation
- 2 Inverse gamma regression model
- 3 Contributions**
- 4 Simulation study
- 5 An application
- 6 Concluding remarks
- 7 References



# Skewness coefficient

The Pearson's standardized third cumulant defined by  $\gamma = \kappa_3 / \kappa_2^{3/2}$ , where  $\kappa_r$  is the  $r$ th cumulant of the distribution, is the most well-known measure of asymmetry.

When  $\gamma > 0$  ( $\gamma < 0$ ) the distribution is positively (negatively) skewed and will have a longer (shorter) right tail and a shorter (longer) left tail.



# Skewness coefficient

The Pearson's standardized third cumulant defined by  $\gamma = \kappa_3 / \kappa_2^{3/2}$ , where  $\kappa_r$  is the  $r$ th cumulant of the distribution, is the most well-known measure of asymmetry.

When  $\gamma > 0$  ( $\gamma < 0$ ) the distribution is positively (negatively) skewed and will have a longer (shorter) right tail and a shorter (longer) left tail.



# Skewness coefficient

There is no closed-form for the skewness coefficient of the distribution of maximum likelihood estimators of the linear parameters in the RIG. The alternative is to derive an approximation up to order  $n^{-1/2}$  from the general expression of Bowman and Shenton (1998).



# Skewness coefficient

After a tedious algebra, we express the third cumulant of the respective set of MLE regressors parameters as,

$$\begin{aligned}\gamma_1(\hat{\beta}) &= \kappa_3(\hat{\beta}) \odot \left\{ \text{diag} \left( \hat{\mathbf{K}}^{\beta, \beta} \right) \right\}^{-3/2}, \\ \gamma_1(\hat{\lambda}) &= \kappa_3(\hat{\lambda}) \odot \left\{ \text{diag} \left( \hat{\mathbf{K}}^{\lambda, \lambda} \right) \right\}^{-3/2},\end{aligned}\tag{2}$$



# Skewness coefficient

where

$$\begin{aligned}\kappa_3(\hat{\beta}) &= \mathbf{Q}_{\beta\beta}^{(3)} \mathbf{M}_1 \mathbf{1} + \left( \mathbf{Q}_{\beta\beta}^{(2)} \odot \mathbf{Q}_{\beta\lambda} \right) (\mathbf{M}_2 + 2\mathbf{M}_3) \mathbf{1} \\ &\quad + \left( \mathbf{Q}_{\beta\beta} \odot \mathbf{Q}_{\beta\lambda}^{(2)} \right) (2\mathbf{M}_4 + \mathbf{M}_5) \mathbf{1} + \mathbf{Q}_{\beta\lambda}^{(3)} \mathbf{M}_6 \mathbf{1}, \\ \kappa_3(\hat{\lambda}) &= \mathbf{Q}_{\lambda\beta}^{(3)} \mathbf{M}_1 \mathbf{1} + \left( \mathbf{Q}_{\lambda\beta}^{(2)} \odot \mathbf{Q}_{\lambda\lambda} \right) (\mathbf{M}_2 + 2\mathbf{M}_3) \mathbf{1} \\ &\quad + \left( \mathbf{Q}_{\lambda\beta} \odot \mathbf{Q}_{\lambda\lambda}^{(2)} \right) (2\mathbf{M}_4 + \mathbf{M}_5) \mathbf{1} + \mathbf{Q}_{\lambda\lambda}^{(3)} \mathbf{M}_6 \mathbf{1}.\end{aligned}\tag{3}$$

The remaining matrices are presented in Magalhães et al. (2025).



# Prediction interval

Prediction intervals are constructed for future realizations of a random variable, whereas confidence intervals are constructed for unknown but fixed model parameters.

Suppose that a prediction for the response of a new observation, denoted by  $Y_0$ , is required. It is natural to assume that  $Y_0 \sim \text{RIG}\left(\exp\{\mathbf{x}_0^\top \boldsymbol{\beta}\}, \exp\{\mathbf{z}_0^\top \boldsymbol{\lambda}\}\right)$ , independently of  $\mathbf{Y}$ .



# Prediction interval

Prediction intervals are constructed for future realizations of a random variable, whereas confidence intervals are constructed for unknown but fixed model parameters.

Suppose that a prediction for the response of a new observation, denoted by  $Y_0$ , is required. It is natural to assume that  $Y_0 \sim \text{RIG}\left(\exp\{\mathbf{x}_0^\top \boldsymbol{\beta}\}, \exp\{\mathbf{z}_0^\top \boldsymbol{\lambda}\}\right)$ , independently of  $\mathbf{Y}$ .



# Prediction interval

After some algebra involving asymptotic theory, omitted here for the sake of brevity, an asymptotic prediction interval for  $Y_0$  is given by,

$$\text{PI}[Y_0, 1 - \alpha] = \left[ \frac{\exp\{\mathbf{x}_0^\top \hat{\beta}\} (1 + \exp\{\mathbf{z}_0^\top \hat{\lambda}\})}{\mathcal{G}_{1-\alpha/2}(\exp\{\mathbf{z}_0^\top \hat{\lambda}\} + 2, 1)}; \frac{\exp\{\mathbf{x}_0^\top \hat{\beta}\} (1 + \exp\{\mathbf{z}_0^\top \hat{\lambda}\})}{\mathcal{G}_{\alpha/2}(\exp\{\mathbf{z}_0^\top \hat{\lambda}\} + 2, 1)} \right], \quad (4)$$



# Prediction interval

After some algebra involving asymptotic theory, omitted here for the sake of brevity, an asymptotic prediction interval for  $Y_0$  is given by,

$$\text{PI}[Y_0, 1 - \alpha] = \left[ \frac{\exp\{\mathbf{x}_0^\top \hat{\beta}\} (1 + \exp\{\mathbf{z}_0^\top \hat{\lambda}\})}{\mathcal{G}_{1-\alpha/2}(\exp\{\mathbf{z}_0^\top \hat{\lambda}\} + 2, 1)}; \frac{\exp\{\mathbf{x}_0^\top \hat{\beta}\} (1 + \exp\{\mathbf{z}_0^\top \hat{\lambda}\})}{\mathcal{G}_{\alpha/2}(\exp\{\mathbf{z}_0^\top \hat{\lambda}\} + 2, 1)} \right], \quad (4)$$

where  $\mathcal{G}_\alpha(a, b)$  denotes the  $\alpha$ th quantile of the  $\mathcal{G}(a, b)$  model. The PI in (4) is simple but yet an effective procedure to compute prediction interval for RIG regression model.



# Prediction interval

After some algebra involving asymptotic theory, omitted here for the sake of brevity, an asymptotic prediction interval for  $Y_0$  is given by,

$$\text{PI}[Y_0, 1 - \alpha] = \left[ \frac{\exp\{\mathbf{x}_0^\top \hat{\beta}\} (1 + \exp\{\mathbf{z}_0^\top \hat{\lambda}\})}{\mathcal{G}_{1-\alpha/2}(\exp\{\mathbf{z}_0^\top \hat{\lambda}\} + 2, 1)}; \frac{\exp\{\mathbf{x}_0^\top \hat{\beta}\} (1 + \exp\{\mathbf{z}_0^\top \hat{\lambda}\})}{\mathcal{G}_{\alpha/2}(\exp\{\mathbf{z}_0^\top \hat{\lambda}\} + 2, 1)} \right], \quad (4)$$

where  $\mathcal{G}_\alpha(a, b)$  denotes the  $\alpha$ th quantile of the  $\mathcal{G}(a, b)$  model. The PI in (4) is simple but yet an effective procedure to compute prediction interval for RIG regression model.



# Outline

- 1 Motivation
- 2 Inverse gamma regression model
- 3 Contributions
- 4 Simulation study**
- 5 An application
- 6 Concluding remarks
- 7 References



# Simulation study

We conducted two experiments.

- ① We compared the agreement between  $\gamma_1$  and the sample skewness;



# Simulation study

We conducted two experiments.

- ① We compared the agreement between  $\gamma_1$  and the sample skewness;
- ② We evaluated the coverage of the prediction interval.



# Simulation study

We conducted two experiments.

- ① We compared the agreement between  $\gamma_1$  and the sample skewness;
- ② We evaluated the coverage of the prediction interval.

The evaluations were performed through a Monte Carlo simulation study with 10,000 replications. All simulations were performed using R (R Core Team, 2024).



# Simulation study

We conducted two experiments.

- ① We compared the agreement between  $\gamma_1$  and the sample skewness;
- ② We evaluated the coverage of the prediction interval.

The evaluations were performed through a Monte Carlo simulation study with 10,000 replications. All simulations were performed using R (R Core Team, 2024).



Table 1: Sample and  $n^{-1/2}$  skewness coefficients.

| $\beta$              | $\lambda$        | Estimator   | $n = 20$ |              | $n = 30$ |              | $n = 40$ |              |
|----------------------|------------------|-------------|----------|--------------|----------|--------------|----------|--------------|
|                      |                  |             | $ \rho $ | $ \gamma_1 $ | $ \rho $ | $ \gamma_1 $ | $ \rho $ | $ \gamma_1 $ |
| (4.50, -0.25, -1.00) | (3.0, -0.1, 0.5) | $\beta_0$   | 0.304    | 0.625        | 0.310    | 0.428        | 0.203    | 0.360        |
|                      |                  | $\beta_1$   | 0.068    | 0.177        | 0.029    | 0.024        | 0.003    | 0.034        |
|                      |                  | $\beta_2$   | 0.071    | 1.769        | 0.104    | 1.482        | 0.155    | 1.037        |
|                      |                  | $\lambda_0$ | 0.854    | 0.416        | 0.786    | 0.241        | 0.924    | 0.191        |
|                      |                  | $\lambda_1$ | 0.011    | 0.326        | 2.249    | 0.237        | 2.775    | 0.214        |
|                      |                  | $\lambda_2$ | 0.481    | 1.331        | 0.027    | 1.066        | 0.509    | 0.787        |

# Simulation study

In general,

- we observed that the sample and the  $n^{-1/2}$  skewness coefficient decrease when the sample size increases;



# Simulation study

In general,

- we observed that the sample and the  $n^{-1/2}$  skewness coefficient decrease when the sample size increases;
- The difference between both coefficients is larger for small sample sizes.



# Simulation study

In general,

- we observed that the sample and the  $n^{-1/2}$  skewness coefficient decrease when the sample size increases;
- The difference between both coefficients is larger for small sample sizes.

For fixed dataset ( $n$ ), the addition of parameters may weaken the approximation of the MLE distribution by normal distribution.



# Simulation study

In general,

- we observed that the sample and the  $n^{-1/2}$  skewness coefficient decrease when the sample size increases;
- The difference between both coefficients is larger for small sample sizes.

For fixed dataset ( $n$ ), the addition of parameters may weaken the approximation of the MLE distribution by normal distribution.



Table 2: Comparison for PI.

| Case | Set | Scenario | n=20  |       |       | n=30  |       |       | n=40  |       |       | n=60  |       |       | n=100 |       |       |
|------|-----|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|      |     |          | 90%   | 95%   | 99%   | 90%   | 95%   | 99%   | 90%   | 95%   | 99%   | 90%   | 95%   | 99%   | 90%   | 95%   | 99%   |
| A    | 1   | I        | 0.768 | 0.830 | 0.909 | 0.831 | 0.892 | 0.957 | 0.858 | 0.916 | 0.971 | 0.869 | 0.924 | 0.977 | 0.881 | 0.934 | 0.983 |
|      |     | II       | 0.836 | 0.895 | 0.957 | 0.862 | 0.919 | 0.976 | 0.872 | 0.927 | 0.977 | 0.875 | 0.934 | 0.987 | 0.893 | 0.941 | 0.987 |
|      |     | III      | 0.855 | 0.913 | 0.969 | 0.866 | 0.923 | 0.977 | 0.875 | 0.932 | 0.980 | 0.890 | 0.942 | 0.987 | 0.888 | 0.940 | 0.984 |
|      | 2   | I        | 0.762 | 0.823 | 0.903 | 0.830 | 0.888 | 0.953 | 0.852 | 0.913 | 0.971 | 0.864 | 0.922 | 0.975 | 0.882 | 0.933 | 0.983 |
|      |     | II       | 0.834 | 0.899 | 0.962 | 0.861 | 0.920 | 0.976 | 0.877 | 0.932 | 0.982 | 0.880 | 0.937 | 0.986 | 0.885 | 0.942 | 0.987 |
|      |     | III      | 0.855 | 0.914 | 0.970 | 0.869 | 0.924 | 0.976 | 0.877 | 0.932 | 0.981 | 0.886 | 0.940 | 0.985 | 0.891 | 0.943 | 0.987 |
| B    | 1   | I        | 0.794 | 0.858 | 0.934 | 0.841 | 0.901 | 0.960 | 0.853 | 0.912 | 0.972 | 0.874 | 0.929 | 0.980 | 0.883 | 0.940 | 0.986 |
|      |     | II       | 0.838 | 0.897 | 0.962 | 0.860 | 0.917 | 0.978 | 0.868 | 0.924 | 0.976 | 0.882 | 0.938 | 0.986 | 0.886 | 0.938 | 0.985 |
|      |     | III      | 0.845 | 0.909 | 0.969 | 0.864 | 0.923 | 0.979 | 0.877 | 0.933 | 0.982 | 0.881 | 0.938 | 0.984 | 0.888 | 0.944 | 0.988 |
|      | 2   | I        | 0.806 | 0.870 | 0.938 | 0.844 | 0.906 | 0.965 | 0.855 | 0.909 | 0.969 | 0.873 | 0.931 | 0.979 | 0.890 | 0.942 | 0.987 |
|      |     | II       | 0.838 | 0.900 | 0.968 | 0.865 | 0.919 | 0.976 | 0.872 | 0.928 | 0.980 | 0.880 | 0.934 | 0.984 | 0.891 | 0.941 | 0.988 |
|      |     | III      | 0.849 | 0.910 | 0.971 | 0.863 | 0.923 | 0.982 | 0.877 | 0.931 | 0.984 | 0.881 | 0.936 | 0.984 | 0.891 | 0.942 | 0.989 |

# Simulation study

- For small sample sizes, the prediction intervals exhibit coverage below their nominal levels;
- Under-coverage is more pronounced in scenarios with a larger number of parameters;



# Simulation study

- For small sample sizes, the prediction intervals exhibit coverage below their nominal levels;
- Under-coverage is more pronounced in scenarios with a larger number of parameters;
- As the sample size increases, the coverage approaches the nominal level.



# Simulation study

- For small sample sizes, the prediction intervals exhibit coverage below their nominal levels;
- Under-coverage is more pronounced in scenarios with a larger number of parameters;
- As the sample size increases, the coverage approaches the nominal level.



# Outline

- 1 Motivation
- 2 Inverse gamma regression model
- 3 Contributions
- 4 Simulation study
- 5 An application**
- 6 Concluding remarks
- 7 References



# Leptinaria unilamellata dataset

From Daniel (2013), we illustrate the proposed method by analyzing a data set in which the response variable is the intraoviductal retention duration in individuals of the species *Leptinaria unilamellata* (d'Orbigny, 1835,  $y$ ),



# Leptinaria unilamellata dataset

with covariates representing environmental conditions that may affect the duration of intraoviductal retention ( $x_2$ ) and the number of offspring released ( $x_3$ ) by each individual was recorded by direct visual counting (Arct et al., 2025; Seki et al., 2010; Setiaji et al., 2021).



# Leptinaria unilamellata dataset

with covariates representing environmental conditions that may affect the duration of intraoviductal retention ( $x_2$ ) and the number of offspring released ( $x_3$ ) by each individual was recorded by direct visual counting (Arct et al., 2025; Seki et al., 2010; Setiaji et al., 2021).



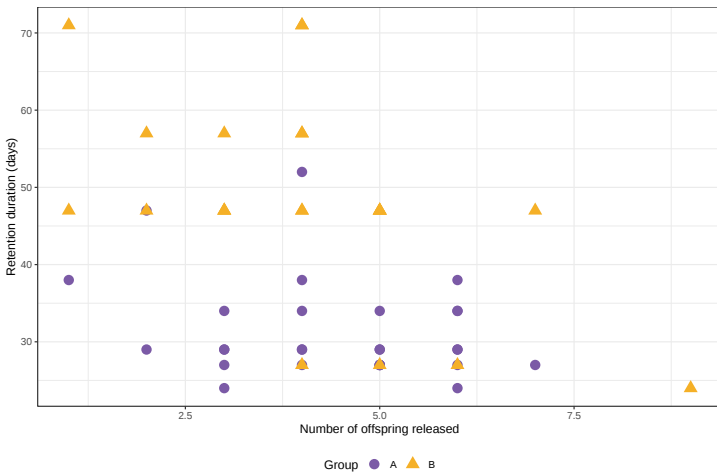


Figure 1: Scatterplot.

# Leptinaria unilamellata dataset

We assumed that  $y_i \sim \text{RIG}(\mu_i, \phi_i)$  with

$$\log \mu_i = \beta_1 + \beta_2 \times \text{Group:B}_i + \beta_3 \times \text{Off}_i$$

$$\log \phi_i = \lambda_1 + \lambda_2 \times \text{Group:B}_i,$$

$$i = 1, \dots, 52.$$



Table 3: Estimates for the *Leptinaria unilamellata* dataset.

| Parameter   | Estimate | SE    | z-value | p-value | $\gamma_1$ |
|-------------|----------|-------|---------|---------|------------|
| $\beta_1$   | 3.683    | 0.084 | 43.683  | < 0.001 | 0.146      |
| $\beta_2$   | 0.410    | 0.063 | 6.501   | < 0.001 | 0.381      |
| $\beta_3$   | -0.055   | 0.018 | -3.094  | 0.003   | 0.000      |
| $\lambda_1$ | 3.482    | 0.269 | 12.928  | < 0.001 | 0.104      |
| $\lambda_2$ | -0.932   | 0.414 | -2.251  | 0.029   | -0.047     |

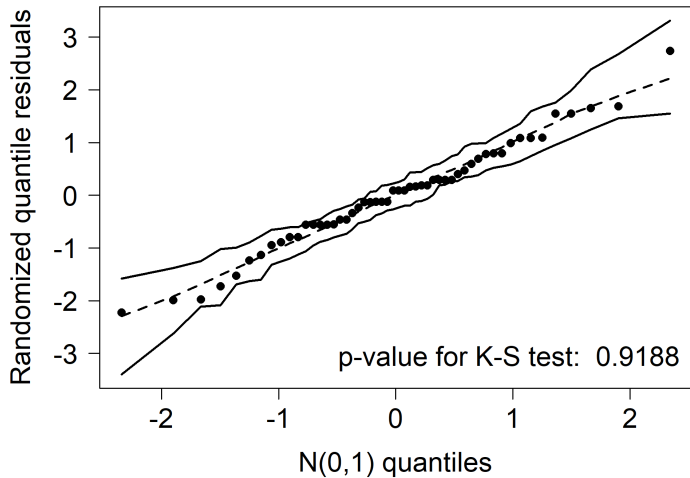


Figure 2: QQ plot for quantile residuals.

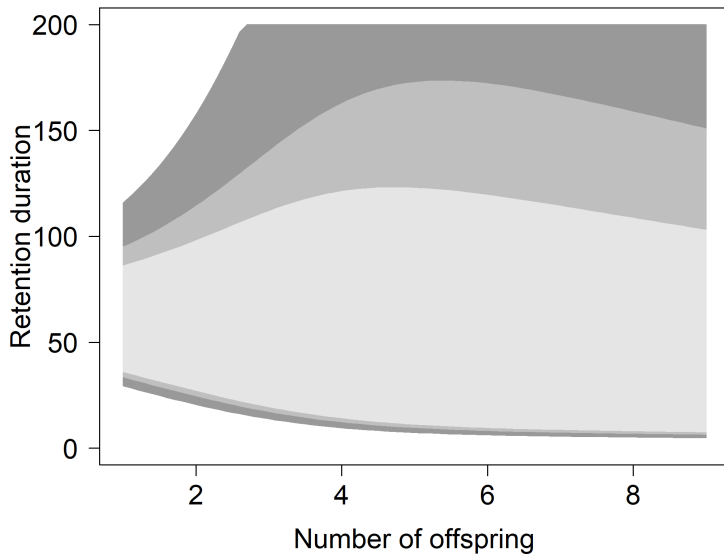


Figure 3: 90%, 95% and 99% PI of the retention duration.

# Outline

- 1 Motivation
- 2 Inverse gamma regression model
- 3 Contributions
- 4 Simulation study
- 5 An application
- 6 Concluding remarks**
- 7 References



# Concluding remarks

In this work, in the Reparameterized Gamma Inverse regression model, we developed a

- ① skewness coefficient, which identifies the shape of the asymptotic distribution of the MLE for regressors parameters;



# Concluding remarks

In this work, in the Reparameterized Gamma Inverse regression model, we developed a

- ① skewness coefficient, which identifies the shape of the asymptotic distribution of the MLE for regressors parameters;
- ② prediction interval that can be narrower, providing more accurate predictions.



# Concluding remarks

In this work, in the Reparameterized Gamma Inverse regression model, we developed a

- ① skewness coefficient, which identifies the shape of the asymptotic distribution of the MLE for regressors parameters;
- ② prediction interval that can be narrower, providing more accurate predictions.



# Acknowledgment

- ① FAPEMIG/Brazil (grants APQ-03202-26 & RED-00133-21);
- ② Graduate Program in Mathematics - UFJF.



# Concluding remarks

Magalhães, T. M., Gallardo, D. I., and Diniz, M. A. (2025). Inferential and predictive procedures for inverse gamma regression model. *Journal of Statistical Computation and Simulation* 95 (11), 2327–2342.



# Outline

- 1 Motivation
- 2 Inverse gamma regression model
- 3 Contributions
- 4 Simulation study
- 5 An application
- 6 Concluding remarks
- 7 References



# References

- Arct, A., Skorb, K., Gustafsson, L., Drobniak, S. M., and Martyka, R. (2025). Ecological drivers and fitness consequences of incubation period in *Ficedula albicollis* (collared flycatcher): Insights from a long-term data set. *Ornithology*, 142(4).
- Bourguignon, M. and Gallardo, D. I. (2020). Reparameterized inverse gamma regression models with varying precision. *Statistica Neerlandica*, 74(4):611–627.



# References

- Bowman, K. O. and Shenton, L. R. (1998). Asymptotic skewness and the distribution of maximum likelihood estimators. *Communications in Statistics - Theory and Methods*, 27(11):2743–2760.
- Daniel, P. A. (2013). Aspectos biológicos e histológicos da reprodução de leptinaria unilamellata (d'orbigny, 1835) (mollusca, pulmonata, subulinidae). Master's thesis, Federal University of Juiz de Fora, Juiz de Fora, Brazil.
- d'Orbigny, A. (1835). *Voyage dans l'Amérique méridionale (1826–1833)*, Volume 5: *Mollusques*. P. Bertrand, Paris.



# References

- Jørgensen, B. (1997). Proper dispersion models (with discussion). *Brazilian Journal of Probability and Statistics*, 11(2):89–140.
- Lehmann, E. L. and Casella, G. (1998). *Theory of point estimation*. Springer, New York, 2nd edition.
- Magalhães, T. M., Gallardo, D. I., and Bourguignon, M. (2021). Improved point estimation for inverse gamma regression models. *Journal of Statistical Computation and Simulation*, 91(12):2444–2456.



# References

- Magalhães, T. M., Gallardo, D. I., and Diniz, M. A. (2025). Inferential and predictive procedures for inverse gamma regression model. *Journal of Statistical Computation and Simulation*, 95(11):2327–2342.
- Montgomery, D. C., Peck, E. A., and Vining, G. G. (2021). *Introduction to linear regression analysis*. Wiley, New York, 6th edition.
- R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.



# References

- Seki, M., Watanabe, N., Ishii, K., Kinoshita, Y., Aihara, T., Takeiri, S., and Otoi, T. (2010). Influence of parity and litter size on gestation length in beagle dogs. *Canadian Journal of Veterinary Research*, 74(1):78–80.
- Setiaji, A., Lestari, D. A., Sutopo, S., Ondho, Y. S., and Kurnianto, E. (2021). Gestation length and litter size of new zealand white grade rabbit. *Jurnal Sain Peternakan Indonesia*, 16(3):288–291.



# Thank you!

✉ [tiago.magalhaes@ufjf.br](mailto:tiago.magalhaes@ufjf.br)

📁 [ufjf.br/tiago\\_magalhaes](https://ufjf.br/tiago_magalhaes)

