UFSM

CIÊNCIA e NATURA



ISSN 2179-460X

OPEN ACCESS

Ci. e Nat., Santa Maria, v. 48, e83861, 2026• https://doi.org/10.5902/2179460X83861
Submissão: 26/05/2023 • Aprovação: 04/07/2025 • Publicação: 25/09/2026

Estatística

Bartlett's correction in measurement errors models

Correção de *Bartlett* em modelos com erros nas variáveis

Kaíque Ferreira Henrique de Souza [I], **Tatiane Ferreira do Nascimento Melo** [I], **Eder Angelo Milani** [I II], **Tiago Moreira Vargas** *(in memoriam)* [I]

[I]Universidade Federal de Goiás, Goiás, GO, Brasil
[II]Universidade Federal de São Carlos, São Carlos, SP, Brasil

RESUMO

Neste trabalho, abordamos a realização de testes em modelos de regressão linear simples (MRLS) com erros nas variáveis, quando o tamanho da amostra é pequeno. Tais modelos são generalizações dos modelos de regressão clássicos e tem sido aplicados em várias áreas como, por exemplo, na medicina e na química. A estatística da razão de verossimilhanças pode ser utilizada para realizar testes sobre os parâmetros desses modelos. No entanto, em amostras pequenas, a distribuição qui-quadrado não fornece uma boa aproximação para distribuição desta estatística, sob a hipotese nula. Assim, construímos um fator de correção de *Bartlett* para o teste da razão de verossimilhanças que reduz a distorção no tamanho do teste, visto que esta estatística modificada possui distribuição qui-quadrado até um erro de ordem $o(n^{-1})$, onde n é o tamanho da amostra. Uma vantagem da metodologia empregada é que a expressão para a correção de Bartlett envolve somente operações simples de matrizes e vetores. É apresentado um estudo de simulação com a comparação da estatística corrigida e não corrigida. Os resultados numéricos favorecem o teste modificado. Para exemplificar a metodologia proposta, mostramos uma aplicação com dados reais. Adicionalmente, desenvolvemos uma rotina utilizando a linguagem R onde é possível obter a estatística do teste da razão de verossimilhanças corrigida por um fator de correção de *Bartlett* para o MRLS com erros nas variáveis.

Palavras-chave: Correção de *Bartlett*; Distribuição normal bivariada; Modelos com erros nas variáveis; Regressão com amostra pequena

ABSTRACT

In this work, we address testing of simple linear regression models with errors in the variables when the sample size is small. Such models are generalizations of classical regression models and have been applied in various fields such as medicine and chemistry. The likelihood ratio statistic can be used to perform tests on the parameters of these models. However, in small samples, the chi-squared

Artigo publicado por Ciência e Natura sob uma licença CC BY-NC-SA 4.0.

distribution does not provide a acurate approximation for the distribution of this statistic under the null hypothesis. Therefore, we construct a Bartlett correction factor for the likelihood ratio test to reduce the distortion in test size, as this modified statistic has a chi-squared distribution up to an error of order $o(n^{-1})$, where n is the sample size. An advantage of the employed methodology is that the expression for Bartlett's correction involves only simple matrix and vector operations. We present a simulation study comparing the corrected and uncorrected statistics. The numerical results favor the modified test. To illustrate the proposed methodology, we demonstrate an application using real data. Additionally, we develop a routine using the R language where it is possible to obtain the statistic for the likelihood ratio test, corrected by a Bartlett correction factor, for the studied model.

Keywords: Bartlett correction; Bivariate normal distribution; Errors-in-Variables models; Small-Sample regression

1 INTRODUÇÃO

Em problemas práticos existem situações em que a covariável é observada com erro de mensuração. Neste caso, a covariável não é observada diretamente e então, temos os modelos com erros de medição, também chamados de modelos com erros nas variáveis (Fuller, 1987). Estes modelos são generalizações dos modelos de regressão clássicos. Temos três classificações para os modelos com erros de medida, são elas: estrutural, funcional e ultraestrutural. Se assumirmos que as covariáveis não-observadas seguem uma distribuição de probabilidade temos o modelo estrutural, neste caso, temos um problema de identificabilidade, para resolver supõem-se que algumas variâncias ou funcões destas seja conhecidas (Chan & Mak, 1979). Quando as covariáveis são tratadas como parâmetros desconhecidos temos o modelo funcional. O modelo ultraestrutural contempla tanto o modelo estrutural quanto o funcional, veja (Fuller, 1987). Neste artigo, consideramos o modelo estrutural.

Os modelos com erros nas variáveis podem ser aplicados nas áreas: medica (Kulathinal et al., 2002), química (Riu & Rius, 1996), mineração (Zavala et al., 2007), entre outras. Um exemplo prático da utilidade destes modelos é apresentado por Aoki et al. (2001) que realizaram um estudo onde o interesse é comparar a eficácia de dois tipos de escovas de dentes na remoção de placa bacteriana; a covariável é o índice de placa antes da escovação e a variável resposta é o índice de placa após a escovação. Neste caso, é razoável supor que a covariável está sujeita a erros de medição, pois a quantidade de placa bacteriana é avaliada imprecisamente e é determinada de forma semelhante antes e após a escovação.

Quando a amostra é pequena temos que as aproximações assintóticas de primeira ordem nas quais testes estatísticos clássicos usualmente são baseados podem ser pouco precisos. Em geral, esse é o caso do teste da razão de verossimilhanças (Neyman & Pearson, 1928), cuja estatística tem distribuição assintótica qui–quadrado sob a hipótese nula. Uma alternativa para resolver este problema é usar teoria assintótica de alta ordem, que pode tornar métodos inferenciais tradicionais mais precisos em pequenas amostras. A estatística da razão de verossimilhanças, por exemplo, pode ser transformada para que tenha distribuição, sob a hipótese nula, mais próxima da distribuição de referência do que a estatística não transformada. Bartlett (1937, 1947, 1954) obteve correções para a estatística da razão de verossimilhanças. Essa nova estatística admite um fator de correção simples, que não depende do valor da estatística original. Em problemas regulares, a estatística da razão de verossimilhanças tem distribuição qui-quadrado a menos de um erro de ordem $o(n^{-1/2})$ sob a hipótese nula, enquanto que a distribuição de sua versão corrigida é qui-quadrado com erro de ordem $o(n^{-1})$.

Mais detalhes sobre correções de Bartlett podem ser encontrados em Cordeiro & Cribari-Neto (1996), que fornecem uma base teórica abrangente sobre o tema. Aplicações dessas correções em diferentes contextos têm sido exploradas em diversos trabalhos. Por exemplo, Aubin & Cordeiro (2000) investigam a aplicação de correções de Bartlett em modelos lineares com matriz de covariância dos erros não escalar; Wong (1991) considera testes com múltiplos coeficientes angulares; Bayer & Cribari-Neto (2013) aplicam a correção no contexto de modelos de regressão beta; Noma et al. (2018) propõem ajustes em metanálises em rede; e Melo et al. (2022) desenvolvem refinamentos de ordem superior em modelos de regressão multivariada do tipo Dirichlet. Trabalhos mais recentes continuam a explorar o uso da correção de Bartlett em contextos diversos. Wong et al. (2025) propõem uma correção do tipo Bartlett para testes de homogeneidade de médias sob a distribuição Normal Inversa; Magalhães et al. (2024) desenvolvem ajustes em modelos lineares generalizados com zero-ajustamento; Araújo et al. (2022) analisam correções de Bartlett e do tipo Bartlett em modelos de regressão não linear heterocedásticos com distribuições simétricas. Esses estudos reforçam a atualidade e a relevância da metodologia em diferentes áreas da estatística aplicada. Tais contribuições evidenciam a ampla aplicabilidade das

correções de Bartlett em distintas classes de modelos, o que destaca ainda mais a importância do presente trabalho ao aplicar essa abordagem em modelos com erros nas variáveis.

A correção de Bartlett tem sido amplamente utilizada para melhorar a aproximação assintótica da estatística da razão de verossimilhanças, especialmente em amostras pequenas. Recentemente, Melo et al. (2024) apresentaram uma formulação geral da correção de Bartlett e do ajuste de Skovgaard em modelos de regressão multivariada com parametrização geral, englobando, como casos particulares, modelos com erros nas variáveis. O presente trabalho se insere nesse contexto ao tratar especificamente do modelo de regressão linear simples com erro nas variáveis, sob duas condições distintas de identificabilidade. A principal contribuição está na dedução explícita das expressões envolvidas na correção de Bartlett para este modelo específico, com o uso de notação matricial, bem como na implementação computacional da metodologia em linguagem R, com código disponibilizado publicamente. Tal abordagem facilita sua aplicação prática e amplia o acesso de pesquisadores ao método, destacando sua relevância mesmo em modelos aparentemente simples, mas que ainda carecem de rotinas prontas para uso.

O objetivo deste trabalho é obter a correção de *Bartlett* para modelos de regressão linear simples com erro nas variáveis (com uma covariável), considerando diferentes condições de identificabilidade. Além disso, desenvolvemos uma rotina em R (R Development Core Team, 2022) que permite ajustar esse tipo de modelo e aplicar a correção de *Bartlett* à estatística da razão de verossimilhanças, proporcionando inferências mais precisas. Estudos de simulação são conduzidos para comparar o desempenho dos testes tradicional e corrigido. Adicionalmente, utilizamos um conjunto de dados reais para ilustrar a aplicação prática da metodologia proposta.

Neste trabalho, utiliza-se a correção de Bartlett aplicada à estatística da razão de verossimilhanças com o objetivo de melhorar a qualidade inferencial em situações com tamanho amostral reduzido. Em tais contextos, a aproximação assintótica da estatística tradicional à distribuição qui-quadrado pode ser imprecisa, levando a decisões incorretas. A correção de Bartlett reduz essa distorção, oferecendo maior aderência à distribuição de referência sob a hipótese nula. Além disso, sua aplicação ao modelo com erros nas variáveis é analiticamente viável e não compromete a estrutura do teste, o que favorece sua adoção em práticas estatísticas.

Este artigo está estruturado da seguinte forma: na Seção 2, apresentamos o modelo de regressão com erros nas variáveis, a correção de *Bartlett* para a estatística da razão de verossimilhanças é apresentado na Seção 3. Um estudo de simulação é apresentado na Seção 4. Uma aplicação a dados reais é apresentado na Seção 5. Na Seção 6 é apresentado uma rotina em linguagem R Development Core Team (2022) para o cálculo das estatísticas da razão de verossimilhança corrigida por *Bartlett*. Na Seção 7, as conclusões do trabalho são apresentadas.

2 MODELOS COM ERROS NAS VARIÁVEIS

Segundo Fuller (1987), o modelo de regressão linear simples é definido por

$$Y_i = \alpha + \beta x_i + e_i,\ i = 1, 2, \ldots, n,$$

adotando que a quantidade x_i não é observada diretamente, temos então o modelo linear com erros nas variáveis. Denotaremos tal erro por u_i, então o valor que se observa é $X_i = x_i + u_i,\ i = 1, 2, \ldots, n$, sendo que

(i) $\mathrm{E}(e_i) = 0$, $\mathrm{Var}(e_i) = \sigma_e^2$ e $\mathrm{Cov}(e_i, e_j) = 0$, $i \neq j$;

(ii) $\mathrm{E}(u_i) = 0$, $\mathrm{Var}(u_i) = \sigma_u^2$ e $\mathrm{Cov}(u_i, u_j) = 0$, $i \neq j$;

(iii) $\mathrm{E}(x_i) = \mu_x$ e $\mathrm{Var}(x_i) = \sigma_x^2$;

(iv) $\mathrm{Cov}(e_i, u_i) = \mathrm{Cov}(x_i, e_i) = \mathrm{Cov}(x_i, u_i) = 0$.

O modelo também pode ser escrito na forma matricial, sendo dado por

$$\boldsymbol{Z}_i = \boldsymbol{\delta} + \Delta \boldsymbol{b}_i,\ i = 1, 2, \ldots, n, \tag{1}$$

com

$$\boldsymbol{Z}_i = \begin{pmatrix} Y_i \\ X_i \end{pmatrix}, \quad \boldsymbol{\delta} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}, \quad \Delta = \begin{pmatrix} \beta & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{b}_i = \begin{pmatrix} x_i \\ e_i \\ u_i \end{pmatrix}.$$

Assumindo que os vetores $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n$ são independentes e que cada $\boldsymbol{b}_i$ tem distribuição normal trivariada com vetor de médias $\boldsymbol{\eta}$ e matriz de covariâncias Ω ($\boldsymbol{b}_i \sim N_3(\boldsymbol{\eta}, \Omega)$), com

$$\boldsymbol{\eta} = \begin{pmatrix} \mu_x \\ 0 \\ 0 \end{pmatrix} \quad \text{e} \quad \Omega = \begin{pmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_e^2 & 0 \\ 0 & 0 & \sigma_u^2 \end{pmatrix},$$

o vetor aleatório $\boldsymbol{Z}_i$ tem distribuição normal bivariada com vetor de médias $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta})$ e matriz de covariâncias $\Sigma = \Sigma(\boldsymbol{\theta})$ positiva definida, em que $\boldsymbol{\theta} = (\beta, \alpha, \mu_x, \sigma_x^2, \sigma_u^2, \sigma_e^2)^\top$, ou seja, $\boldsymbol{Z}_i \sim N_2(\boldsymbol{\mu}, \Sigma)$, com $\boldsymbol{\mu} = \boldsymbol{\delta} + \Delta\boldsymbol{\eta}$ e $\Sigma = \Delta\Omega\Delta^\top$. Matricialmente, podemos escrever $\boldsymbol{\mu}$ e Σ da forma

$$\boldsymbol{\mu} = \begin{pmatrix} \alpha + \beta\mu_x \\ \mu_x \end{pmatrix} \quad \text{e} \quad \Sigma = \begin{pmatrix} \beta^2\sigma_x^2 + \sigma_e^2 & \beta\sigma_x^2 \\ \beta\sigma_x^2 & \sigma_x^2 + \sigma_u^2 \end{pmatrix}.$$

O modelo estrutural definido em (1) não é identificável, pois a relação entre o vetor de parâmetros da distribuição de $\boldsymbol{Z}_i$ e o vetor $(\beta, \alpha, \mu_x, \sigma_x^2, \sigma_u^2, \sigma_e^2)$ não é um-a-um, para mais detalhes ver Fuller (1987). A fim de contornar o problema de não-identificabilidade, consideram-se suposições sobre as variâncias σ_x^2 e σ_e^2. É comum supor que a razão $\lambda_e = \sigma_e^2/\sigma_u^2$ ou $\lambda_x = \sigma_x^2/\sigma_u^2$ é conhecida, veja em Cheng & Van Ness (1999). Aqui, consideramos inferências sob as duas suposições, isto é, admitimos, separadamente, que as razões λ_e e λ_x são conhecidas, tendo como consequência modificações no vetor $\boldsymbol{\theta}$ e na matriz Σ, que são dadas por:

(i) se λ_e é conhecida, então
$$\begin{cases} \boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)^\top = (\beta, \alpha, \mu_x, \sigma_x^2, \sigma_u^2)^\top, \\ \Sigma = \begin{pmatrix} \beta^2\sigma_x^2 + \lambda_e\sigma_u^2 & \beta\sigma_x^2 \\ \beta\sigma_x^2 & \sigma_x^2 + \sigma_u^2 \end{pmatrix}, \end{cases}$$

(ii) se λ_x é conhecida, então
$$\begin{cases} \boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)^\top = (\beta, \alpha, \mu_x, \sigma_u^2, \sigma_e^2)^\top, \\ \Sigma = \begin{pmatrix} \lambda_x\beta^2\sigma_u^2 + \sigma_e^2 & \lambda_x\beta\sigma_u^2 \\ \lambda_x\beta\sigma_u^2 & (\lambda_x + 1)\sigma_u^2 \end{pmatrix}. \end{cases}$$

Considerando uma amostra de tamanho n, onde é observado para cada unidade os valores Y_i e X_i, as estimativas de máxima verossimilhança são obtidas via logaritmo da função de verossimilhança para o modelo, que é dada por

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2}\log|\Sigma| - \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{z}_i - \boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{z}_i - \boldsymbol{\mu}),$$

sendo $\boldsymbol{z}_i = (Y_i, X_i)^\top$. Para estimação dos parâmetros do modelo com erros nas variáveis podemos usar os métodos de mínimos quadrados, momentos e máxima verossimilhança. Sob a suposição de normalidade os estimadores de método dos momentos são os estimadores de máxima verossimilhança ajustados para graus de liberdade (Fuller, 1987). Os estimadores de momentos são dados por: $\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X}$ e $\widehat{\mu_x} = \overline{X}$. Quando λ_e é conhecida temos:

$$\widehat{\beta} = \frac{S_{YY} - \lambda_e S_{XX} + [(S_{YY} - \lambda_e S_{XX})^2 + 4\lambda_e S_{XY}^2]^{1/2}}{2S_{XY}},\ \widehat{\sigma}_x^2 = \frac{S_{XY}}{\widehat{\beta}} \text{ e } \widehat{\sigma}_u^2 = \frac{S_{YY} - 2\widehat{\beta}S_{XY} + \widehat{\beta}^2 S_{XX}}{\widehat{\beta}^2 + \lambda_e}.$$

Se λ_x é conhecida:

$$\widehat{\beta} = \frac{S_{XY}(\lambda_x + 1)}{\lambda_x S_{XX}},\ \widehat{\sigma}_u^2 = \frac{S_{XX}}{\lambda_x + 1} \text{ e } \widehat{\sigma}_e^2 = \frac{\lambda_x(S_{YY}S_{XX} - S_{XY}^2) - S_{XY}^2}{\lambda_x S_{XX}}.$$

3 Correção de *Bartlett* para o modelo com erros nas variáveis

O vetor de parâmetros $\boldsymbol{\theta}$ é particionado como $\boldsymbol{\theta} = (\psi, \boldsymbol{\omega}^\top)^\top$, sendo ψ o parâmetro de interesse (escalar) e $\boldsymbol{\omega}$ o vetor de parâmetros de perturbação. Neste trabalho nosso interesse consiste em testar hipóteses da forma: $\mathcal{H}_0 : \psi = \psi^{(0)} \ \text{ versus } \ \mathcal{H}_1 : \psi \neq \psi^{(0)}$, em que $\psi^{(0)}$ é uma constante. Aqui, consideramos $\psi = \theta_1 = \beta$ e $\boldsymbol{\omega} = (\theta_2, \theta_3, \theta_4, \theta_5)^\top$. O estimador de máxima verossimilhança irrestrito de $\boldsymbol{\theta}$ é denotado por $\widehat{\boldsymbol{\theta}} = (\widehat{\psi}, \widehat{\boldsymbol{\omega}}^\top)^\top$ e o correspondente estimador sob a hipótese nula por $\widetilde{\boldsymbol{\theta}} = (\widetilde{\psi}^{(0)}, \widetilde{\boldsymbol{\omega}}^\top)^\top$. Usamos "$\widehat{\ }$" em matrizes e vetores para indicar que os mesmos são avaliados nas estimativas de máxima verossimilhança irrestritas e "$\widetilde{\ }$" quando são avaliados nas estimativas de máxima verossimilhança sob a hipótese nula, por exemplo, $\widehat{\boldsymbol{\mu}} = \boldsymbol{\mu}(\widehat{\boldsymbol{\theta}})$, $\widetilde{\boldsymbol{\mu}} = \boldsymbol{\mu}(\widetilde{\boldsymbol{\theta}})$, $\widehat{\Sigma} = \Sigma(\widehat{\boldsymbol{\theta}})$ e $\widetilde{\Sigma} = \Sigma(\widetilde{\boldsymbol{\theta}})$.

A estatística da razão de verossimilhanças usual para o teste de $\mathcal{H}_0$ baseada na função de verossimilhança é dada por

$$LR = 2\left(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})\right),$$

em problemas regulares, está estatística converge para a distribuição χ^2_1. Entretanto, em pequenas amostras, a aproximação de primeira ordem pode não ser satisfatória. Com o objetivo de melhorar esta aproximação, Bartlett (1937, 1947, 1954) propôs multiplicar a estatística LR por uma constante, C^{-1}, obtendo assim uma nova estatística, conhecida como estatística da razão de verossimilhanças corrigida por *Bartlett*:

$$LR_{B_1} = \frac{LR}{C}.$$

Fatores de correção de Bartlett para a estatística da razão de verossimilhanças também foram desenvolvidos em problemas regulares, esta nova estatística também tem distribuição χ^2_1 a menos de um erro de ordem n^{-2}, sob $\mathcal{H}_0$, isto é, $P(LR_{B_1} > x_\gamma) = \gamma + o(n^{-1})$, enquanto que $P(LR > x_\gamma) = \gamma + o(n^{-1/2})$, em que x_γ é o quantil de ordem $(1-\gamma)$ da distribuição χ^2_1; veja em Barndorff-Nielsen (1986). As seguintes estatísticas de testes corrigidas por *Bartlett* são equivalentes e fornecem a mesma ordem de precisão de LR_{B_1} (Cordeiro & Cribari-Neto, 2014), sendo elas:

$$\begin{aligned} LR_{B_2} &= LR\exp\{1-C\}, \\ LR_{B_3} &= LR(2-C). \end{aligned}$$

Uma expressão matricial para C em termos de cumulantes do logaritmo da função de verossimilhança foi desenvolvida por Cordeiro (1993). Esta expressão é dada por:

$$C = 1 + \epsilon_{(p)} - \epsilon_{(p-1)}, \tag{2}$$

onde $\epsilon_{(p)} = \operatorname{tr}[K^{-1}(L - M - N)]$, com K sendo a matriz de informação de Fisher, a inversa da matriz de informação de Fisher é denotada por K^{-1}. Além disso, os (r,s)-ésimos elementos das matrizes L, M e N são, respectivamente, K_{rs}, L_{rs}, M_{rs} e N_{rs}, com

$$
\begin{aligned}
L_{rs} &= \operatorname{tr}\left(K^{-1}A^{(rs)}\right), \\
M_{rs} &= -\frac{1}{6}\operatorname{tr}\left(K^{-1}P^{(r)}K^{-1}P^{(s)}\right) + \operatorname{tr}\left(K^{-1}P^{(r)}K^{-1}[Q^{(s)}]^{\top}\right) - \operatorname{tr}\left(K^{-1}Q^{(r)}K^{-1}Q^{(s)}\right), \\
N_{rs} &= -\frac{1}{4}\operatorname{tr}\left(P^{(r)}K^{-1}\right)\operatorname{tr}\left(P^{(s)}K^{-1}\right) + \operatorname{tr}\left(P^{(r)}K^{-1}\right)\operatorname{tr}\left(Q^{(s)}K^{-1}\right) - \operatorname{tr}\left(Q^{(r)}K^{-1}\right) \\
&\quad \operatorname{tr}\left(Q^{(s)}K^{-1}\right).
\end{aligned}
$$

Para o modelo com erros nas variáveis, definido por (1), temos que os (t, w)-ésimos elementos das matrizes $P^{(r)}$, $Q^{(s)}$ e $A^{(rs)}$ são, respectivamente, $P_{tw}^{(r)}$, $Q_{tw}^{(s)}$ e $A_{tw}^{(rs)}$, com

$$
\begin{aligned}
P_{tw}^{(r)} &= -\frac{n}{2}\operatorname{tr}\left[\Sigma^{\theta_t\theta_r}\Sigma_{\theta_w} + \Sigma^{\theta_r}\Sigma_{\theta_t\theta_w} + \Sigma^{\theta_t}\Sigma_{\theta_r\theta_w} + \Sigma^{-1}\Sigma_{\theta_t\theta_r\theta_w} + \Sigma^{\theta_t\theta_r\theta_w}\Sigma\right] \\
&\quad -n\{\mu_{\theta_t}^{\top}\Sigma^{\theta_w}\mu_{\theta_r} + \mu_{\theta_r}^{\top}\Sigma^{\theta_t}\mu_{\theta_w} + \mu_{\theta_t}^{\top}\Sigma^{\theta_r}\mu_{\theta_w} + \mu_{\theta_t}^{\top}\Sigma^{-1}\mu_{\theta_r\theta_w} + \mu_{\theta_r}^{\top}\Sigma^{-1}\mu_{\theta_t\theta_w} \\
&\quad +\mu_{\theta_w}^{\top}\Sigma^{-1}\mu_{\theta_t\theta_r}\},
\end{aligned}
$$

$$
\begin{aligned}
Q_{tw}^{(s)} &= -\frac{n}{2}\operatorname{tr}\left[\Sigma^{\theta_t\theta_s}\Sigma_{\theta_w} + \Sigma^{\theta_s\theta_w}\Sigma_{\theta_t} + \Sigma^{\theta_s}\Sigma_{\theta_t\theta_w} + \Sigma^{\theta_t}\Sigma_{\theta_s\theta_w} + \Sigma^{-1}\Sigma_{\theta_t\theta_s\theta_w} + \Sigma^{\theta_t\theta_s\theta_w}\Sigma\right] \\
&\quad -n\{\mu_{\theta_s}^{\top}\Sigma^{\theta_t}\mu_{\theta_w} + \mu_{\theta_w}^{\top}\Sigma^{-1}\mu_{\theta_t\theta_s} + \mu_{\theta_s}^{\top}\Sigma^{-1}\mu_{\theta_t\theta_w}\},
\end{aligned}
$$

$$
\begin{aligned}
A_{tw}^{(rs)} &= -\frac{n}{8}tr\big(\Sigma^{\theta_r\theta_w\theta_t}\Sigma_{\theta_s} + 4\Sigma^{\theta_s\theta_r\theta_t}\Sigma_{\theta_w} + \Sigma^{\theta_r}\Sigma_{\theta_w\theta_t\theta_s} + \Sigma^{\theta_w}\Sigma_{\theta_r\theta_t\theta_s} + \Sigma^{\theta_t}\Sigma_{\theta_r\theta_w\theta_s} \\
&\qquad +\Sigma^{\theta_r\theta_w}\Sigma_{\theta_t\theta_s} + \Sigma^{\theta_w\theta_t}\Sigma_{\theta_r\theta_s} + 5\Sigma^{\theta_r\theta_t}\Sigma_{\theta_w\theta_s} + \Sigma^{\theta_r\theta_w\theta_t\theta_s}\Sigma\big) \\
&\quad -\frac{n}{4}\Big[\boldsymbol{\mu}_{\theta_r}^{\top}\Sigma^{\theta_w\theta_s}\boldsymbol{\mu}_{\theta_t} + \boldsymbol{\mu}_{\theta_r}^{\top}\Sigma^{\theta_w\theta_t}\boldsymbol{\mu}_{\theta_s} - 3\boldsymbol{\mu}_{\theta_r}^{\top}\Sigma^{\theta_t\theta_s}\boldsymbol{\mu}_{\theta_w} - 3\boldsymbol{\mu}_{\theta_w}^{\top}\Sigma^{\theta_r\theta_s}\boldsymbol{\mu}_{\theta_t} \\
&\quad +\boldsymbol{\mu}_{\theta_s}^{\top}\Sigma^{\theta_r\theta_w}\boldsymbol{\mu}_{\theta_t} + \boldsymbol{\mu}_{\theta_s}^{\top}\Sigma^{\theta_r\theta_t}\boldsymbol{\mu}_{\theta_w} + \boldsymbol{\mu}_{\theta_r}^{\top}\Sigma^{\theta_s}\boldsymbol{\mu}_{\theta_w\theta_t} - 3\boldsymbol{\mu}_{\theta_r}^{\top}\Sigma^{\theta_t}\boldsymbol{\mu}_{\theta_w\theta_s} + \boldsymbol{\mu}_{\theta_r}^{\top}\Sigma^{\theta_w}\boldsymbol{\mu}_{\theta_t\theta_s} \\
&\quad +\boldsymbol{\mu}_{\theta_s}^{\top}\Sigma^{\theta_r}\boldsymbol{\mu}_{\theta_w\theta_t} + \boldsymbol{\mu}_{\theta_s}^{\top}\Sigma^{\theta_w}\boldsymbol{\mu}_{\theta_r\theta_t} + \boldsymbol{\mu}_{\theta_s}^{\top}\Sigma^{\theta_t}\boldsymbol{\mu}_{\theta_r\theta_w} + \boldsymbol{\mu}_{\theta_t}^{\top}\Sigma^{\theta_w}\boldsymbol{\mu}_{\theta_r\theta_s} - 3\boldsymbol{\mu}_{\theta_t}^{\top}\Sigma^{\theta_r}\boldsymbol{\mu}_{\theta_w\theta_s} \\
&\quad +\boldsymbol{\mu}_{\theta_t}^{\top}\Sigma^{\theta_s}\boldsymbol{\mu}_{\theta_r\theta_w} - 3\boldsymbol{\mu}_{\theta_w}^{\top}\Sigma^{\theta_s}\boldsymbol{\mu}_{\theta_r\theta_t} - 3\boldsymbol{\mu}_{\theta_w}^{\top}\Sigma^{\theta_t}\boldsymbol{\mu}_{\theta_r\theta_s} - 3\boldsymbol{\mu}_{\theta_w}^{\top}\Sigma^{\theta_r}\boldsymbol{\mu}_{\theta_t\theta_s} \\
&\quad +\boldsymbol{\mu}_{\theta_r\theta_w}^{\top}\Sigma^{-1}\boldsymbol{\mu}_{\theta_t\theta_s} - 3\boldsymbol{\mu}_{\theta_r\theta_t}^{\top}\Sigma^{-1}\boldsymbol{\mu}_{\theta_w\theta_s} + \boldsymbol{\mu}_{\theta_w\theta_t}^{\top}\Sigma^{-1}\boldsymbol{\mu}_{\theta_r\theta_s}\Big]
\end{aligned}
$$

para $r, s, t, w = 1, 2, \ldots, p$, com $p = 5$. Aqui, θ_r representa o r-ésimo elemento do vetor $\boldsymbol{\theta}$, μ_{θ_r} e $\mu_{\theta_r\theta_s}$ são, respectivamente, derivadas de primeira e segunda ordem do vetor $\boldsymbol{\mu}$ com relação aos parâmetros. A mesma notação é usada para as derivadas da matriz

Σ. Temos que Σ_{θ_r}, $\Sigma_{\theta_s\theta_r}$ e $\Sigma_{\theta_t\theta_s\theta_r}$ são derivadas de primeira, segunda e terceira ordem da matriz Σ. Com relação a sua inversa temos que Σ^{θ_r}, $\Sigma^{\theta_s\theta_r}$, $\Sigma^{\theta_t\theta_s\theta_r}$ e $\Sigma^{\theta_w\theta_t\theta_s\theta_r}$ são as respectivas derivadas até quarta ordem de Σ^{-1}. Essas derivadas são definidas a seguir. Aqui, $\Sigma^{\theta_r} = -\Sigma^{-1}\Sigma_{\theta_r}\Sigma^{-1}$ e $\Sigma^{\theta_s\theta_r} = -\Sigma^{\theta_s}\Sigma_{\theta_r}\Sigma^{-1} - \Sigma^{-1}\Sigma_{\theta_s\theta_r}\Sigma^{-1} - \Sigma^{-1}\Sigma_{\theta_r}\Sigma^{\theta_s}$. Além disso,

$$\begin{aligned}\Sigma^{\theta_t\theta_s\theta_r} &= -\Sigma^{\theta_t\theta_s}\Sigma_{\theta_r}\Sigma^{-1} - \Sigma^{\theta_s}\Sigma_{\theta_t\theta_r}\Sigma^{-1} - \Sigma^{\theta_s}\Sigma_{\theta_r}\Sigma^{\theta_t} - \Sigma^{\theta_t}\Sigma_{\theta_s\theta_r}\Sigma^{-1} - \Sigma^{-1}\Sigma_{\theta_t\theta_s\theta_r}\Sigma^{-1} \\ &\quad -\Sigma^{-1}\Sigma_{\theta_s\theta_r}\Sigma^{\theta_t} - \Sigma^{\theta_t}\Sigma_{\theta_r}\Sigma^{\theta_s} - \Sigma^{-1}\Sigma_{\theta_t\theta_r}\Sigma^{\theta_s} - \Sigma^{-1}\Sigma_{\theta_r}\Sigma^{\theta_t\theta_s}\end{aligned}$$

e

$$\begin{aligned}\Sigma^{\theta_w\theta_t\theta_s\theta_r} &= -\Sigma^{\theta_w\theta_t\theta_s}\Sigma_{\theta_r}\Sigma^{-1} - \Sigma^{\theta_t\theta_s}\Sigma_{\theta_w\theta_r}\Sigma^{-1} - \Sigma^{\theta_t\theta_s}\Sigma_{\theta_r}\Sigma^{\theta_w} - \Sigma^{\theta_w\theta_s}\Sigma_{\theta_t\theta_r}\Sigma^{-1} \\ &\quad -\Sigma^{\theta_s}\Sigma_{\theta_w\theta_t\theta_r}\Sigma^{-1} - \Sigma^{\theta_s}\Sigma_{\theta_t\theta_r}\Sigma^{\theta_w} - \Sigma^{\theta_w\theta_s}\Sigma_{\theta_r}\Sigma^{\theta_t} - \Sigma^{\theta_s}\Sigma_{\theta_w\theta_r}\Sigma^{\theta_t} \\ &\quad -\Sigma^{\theta_s}\Sigma_{\theta_r}\Sigma^{\theta_w\theta_t} - \Sigma^{\theta_t}\Sigma_{\theta_w\theta_s\theta_r}\Sigma^{-1} - \Sigma^{\theta_t}\Sigma_{\theta_s\theta_r}\Sigma^{\theta_w} - \Sigma^{\theta_w}\Sigma_{\theta_t\theta_s\theta_r}\Sigma^{-1} \\ &\quad -\Sigma^{-1}\Sigma_{\theta_w\theta_t\theta_s\theta_r}\Sigma^{-1} - \Sigma^{-1}\Sigma_{\theta_t\theta_s\theta_r}\Sigma^{\theta_w} - \Sigma^{\theta_w}\Sigma_{\theta_s\theta_r}\Sigma^{\theta_t} - \Sigma^{-1}\Sigma_{\theta_w\theta_s\theta_r}\Sigma^{\theta_t} \\ &\quad -\Sigma^{-1}\Sigma_{\theta_s\theta_r}\Sigma^{\theta_w\theta_t} - \Sigma^{\theta_w\theta_t}\Sigma_{\theta_r}\Sigma^{\theta_s} - \Sigma^{\theta_t}\Sigma_{\theta_w\theta_r}\Sigma^{\theta_s} - \Sigma^{\theta_t}\Sigma_{\theta_r}\Sigma^{\theta_w\theta_s} \\ &\quad -\Sigma^{\theta_w}\Sigma_{\theta_t\theta_r}\Sigma^{\theta_s} - \Sigma^{-1}\Sigma_{\theta_w\theta_t\theta_r}\Sigma^{\theta_s} - \Sigma^{-1}\Sigma_{\theta_t\theta_r}\Sigma^{\theta_w\theta_s} - \Sigma^{\theta_w}\Sigma_{\theta_r}\Sigma^{\theta_t\theta_s} \\ &\quad -\Sigma^{-1}\Sigma_{\theta_w\theta_r}\Sigma^{\theta_t\theta_s} - \Sigma^{-1}\Sigma_{\theta_r}\Sigma^{\theta_w\theta_t\theta_s}.\end{aligned}$$

Para completar o cálculo de (2) definimos o (r,s)-ésimo elemento da matriz de informação de Fisher K, dado por

$$K_{rs} = \frac{n}{2}tr\left(\Sigma^{\theta_s}\Sigma_{\theta_r}\right) - n\mu_{\theta_r}^{\top}\Sigma^{-1}\mu_{\theta_s}.$$

As derivadas até terceira ordem do vetor $\boldsymbol{\mu}$ e da matriz Σ são nulas, exceto para os casos abaixo.

Condição 1: Quando a razão $\lambda_e = \sigma_e^2/\sigma_u^2$ é conhecida temos os seguintes vetores de $\boldsymbol{\mu}$ que não são nulos: $\mu_{\theta_1} = (1,0)^{\top}$, $\mu_{\theta_2} = (\mu_x,0)^{\top}$, $\mu_{\theta_3} = (\beta,1)^{\top}$, $\mu_{\theta_3\theta_2} = \mu_{\theta_2\theta_3} = (1,0)^{\top}$, e as seguintes matriz Σ que não são nulas

$$\Sigma_{\theta_2} = \begin{pmatrix} 2\beta\sigma_x^2 & \sigma_x^2 \\ \sigma_x^2 & 0 \end{pmatrix}, \Sigma_{\theta_4} = \begin{pmatrix} \beta^2 & \beta \\ \beta & 1 \end{pmatrix}, \Sigma_{\theta_5} = \begin{pmatrix} \lambda_e & 0 \\ 0 & 1 \end{pmatrix},$$

$$\begin{aligned}\Sigma_{\theta_2\theta_2} &= \begin{pmatrix} 2\sigma_x^2 & 0 \\ 0 & 0 \end{pmatrix}, \Sigma_{\theta_4\theta_2} = \Sigma_{\theta_2\theta_4} = \begin{pmatrix} 2\beta & 1 \\ 1 & 0 \end{pmatrix}, \\ \Sigma_{\theta_4\theta_2\theta_2} &= \Sigma_{\theta_2\theta_4\theta_2} = \Sigma_{\theta_2\theta_2\theta_4} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.\end{aligned}$$

Condição 2: Quando a razão $\lambda_x = \sigma_x^2/\sigma_u^2$ é conhecida, as derivadas do vetor $\boldsymbol{\mu}$ são iguais aos dados na Condição 1. Para as derivadas da matriz Σ que não são nulas temos:

$$\begin{aligned}\Sigma_{\theta_2} &= \begin{pmatrix} 2\lambda_x\sigma_u^2\beta & \lambda_x\sigma_u^2 \\ \lambda_x\sigma_u^2 & 0 \end{pmatrix}, \Sigma_{\theta_4} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \Sigma_{\theta_5} = \begin{pmatrix} \beta^2\lambda_x & \beta\lambda_x \\ \beta\lambda_x & \lambda_x + 1 \end{pmatrix}, \\ \Sigma_{\theta_2\theta_2} &= \begin{pmatrix} 2\lambda_x\sigma_u^2 & 0 \\ 0 & 0 \end{pmatrix}, \Sigma_{\theta_5\theta_2} = \Sigma_{\theta_2\theta_5} = \begin{pmatrix} 2\lambda_x\beta & \lambda_x \\ \lambda_x & 0 \end{pmatrix}, \\ \Sigma_{\theta_5\theta_2\theta_2} &= \Sigma_{\theta_2\theta_5\theta_2} = \Sigma_{\theta_2\theta_2\theta_5} = \begin{pmatrix} 2\lambda_x & 0 \\ 0 & 0 \end{pmatrix}.\end{aligned}$$

Cabe destacar que a fórmula geral para o fator de correção de Bartlett utilizada neste trabalho segue o desenvolvimento matricial apresentado em Cordeiro (1993). A novidade do presente estudo consiste na obtenção explícita dessas expressões para o modelo com erro nas variáveis estruturais, sob duas diferentes condições de identificabilidade, e na implementação prática dessa metodologia por meio de uma rotina desenvolvida na linguagem R.

4 ESTUDO DO ERRO TIPO I

Com o intuito de verificar possíveis melhorias nos testes da razão de verossimilhanças corrigidos utilizando a teoria de *Bartlett*, estruturamos um estudo de simulação com o objetivo de estimar a probabilidade do erro tipo I do teste $\mathcal{H}_0$: $\beta = 1$, quando consideramos a estatística com (LR_{B_1}, LR_{B_2} e LR_{B_3}) e sem (LR) correção. Para isso, a geração levou em cosideração as condições sob a hipótese nula e os níveis nominais adotados foram γ = 1%, 5% e 10%.

As simulações foram baseadas em 10.000 (dez mil) repetições de Monte Carlo, considerando amostras de tamanhos $n = 10, 17, 20$ e 50, adotando os valores dos parâmetros como: $\alpha = -5,851$, $\mu_x = 19,824$, $\sigma_x^2 = 14,657$, $\sigma_u^2 = 2,655$ e $\sigma_e^2 = 2,0$. Também

consideramos os dois cenários de acordo com as condições de identificabilidade, sendo a razão λ_e ($\lambda_e = 1$) ou λ_x ($\lambda_x = 1$) conhecida. Para cada cenário e tamanho amostral, foi obtida a proporção de vezes que a hipótese nula foi rejeitada, para cada estatística, os resultados são apresentados na Tabela 1.

Note que o teste baseado na estatística da razão de verossimilhanças original (LR) sempre apresenta taxa de rejeição maior que o nível de significância considerado, isso para todos os tamanhos amostrais adotado no estudo. Os testes da razão de verossimilhanças corrigidos através da correção de *Bartlett* (LR_{B_1}, LR_{B_2} e LR_{B_3}) apresentam melhores desempenhos quando comparado com o teste da razão de verossimilhanças original (LR), para todos as configurações em estudo, pois apresentam taxas de rejeição mais próximas dos níveis nominais. Isso corrobora com a teoria destacada, pois as estatísticas corrigidas apresentam ordem de convergência $o(n^{-1})$, enquanto que a estatística tradicional, a convergência é da ordem $o(n^{-1/2})$.

Quando comparamos os cenários com λ_e ou λ_x conhecido, observamos que as taxas de rejeição parecem não ser influenciadas, pois para todos os tamanhos amostrais e níveis de confiança adotados, os resultados são próximos, não havendo uma grande disparidade nos resultados.

Também é importante destacar que a medida que o tamanho da amostra aumenta, os valores das taxas relacionadas a estatística tradicional (LR) se aproximam dos níveis de significância nominal, como era esperado. Um exemplo, fixando λ_e conhecida e $\gamma = 10\%$, a sequência obtida é $14,96$, $12,18$, $11,76$, $11,48$ e $11,04$, respectivamente para $n = 10,\ 17,\ 20,\ 30$ e $,\ 50$. Enquanto isso, as taxas para as estatísticas corrigidas ocilam ao redor do nível nominal.

Tabela 1 – Taxas de rejeições (em %) da hipótese $\mathcal{H}_0$: $\beta = 1$, para estatísticas com e sem correção (Continua)

		λ_e **é conhecida**				λ_x **é conhecida**			
n	γ	LR	LR_{B_1}	LR_{B_2}	LR_{B_3}	LR	LR_{B_1}	LR_{B_2}	LR_{B_3}
	1%	2,30	1,05	0,94	0,82	2,38	1,22	1,15	1,05
10	5%	8,74	5,40	5,04	4,64	8,82	5,79	5,38	4,81
	10%	14,96	10,44	9,93	9,28	14,90	10,53	10,16	9,63

Tabela 1 – Taxas de rejeições (em %) da hipótese $\mathcal{H}_0$: $\beta = 1$, para estatísticas com e sem correção

		λ_e **é conhecida**				λ_x **é conhecida**			
n	γ	LR	LR_{B_1}	LR_{B_2}	LR_{B_3}	LR	LR_{B_1}	LR_{B_2}	LR_{B_3}
	1%	1,64	0,91	0,87	0,82	1,65	1,01	0,95	0,93
17	5%	6,88	5,02	4,95	4,83	6,56	5,01	4,94	4,87
	10%	12,18	9,68	9,57	9,33	12,35	9,97	9,79	9,60
	1%	1,48	1,04	1,03	1,00	1,39	0,86	0,82	0,79
20	5%	6,32	4,81	4,76	4,66	6,45	4,90	4,78	4,65
	10%	11,76	9,63	9,49	9,36	12,19	9,93	9,82	9,70
	1%	1,28	0,91	0,90	0,90	1,28	0,86	0,81	0,80
30	5%	5,73	4,85	4,80	4,79	5,83	4,85	4,83	4,81
	10%	11,48	10,06	9,99	9,92	11,32	10,01	9,91	9,89
	1%	1,25	1,00	0,99	0,98	1,38	1,16	1,16	1,14
50	5%	5,56	4,92	4,92	4,90	5,55	5,01	5,01	5,00
	10%	11,04	10,13	10,10	10,06	11,30	10,44	10,44	10,38

Fonte: Autores (2023)

5 APLICAÇÃO COM DADOS REAIS

Apresentamos uma aplicação com dados reais para exemplificar a técnica proposta. Os dados foram retirados do Fuller (1987), que por sua fez obteve junto a Comissão de Conservação de *Iowa*. Os dados são relacionados ao número de faisões (galinhas) em *Iowa*, observados duas vezes durante o ano. Esses índices são baseados no número médio de aves avistadas por observadores treinados que viajam por várias rotas específicas no final de abril e início de maio, e novamente em agosto. As observações foram feitas entre os anos de 1962 e 1976, totalizando 15 observações.

O interesse é relacionar os valores obtidos na primavera com os valores de agosto, no entanto não se espera que a relação seja uma função linear perfeita, mas se almeja obter uma relação razoável, sendo assim, adotamos o modelo descrito na equação (1), sendo que a variável X_i é o número médio de aves avistada durante a primavera e a variável Y_i é o número médio de aves avistada em agosto, com $i = 1, 2, \ldots, 15$. Como destacado em Fuller (1987), ambas as medidas estão sujeitas a

erros por dois motivos, sendo eles: as rotas são uma amostra de todas as rotas possíveis em *lowa*, e não se pode esperar que os observadores avistem todos os faisões ao longo da rota. Com isso justificamos a utilização do modelo com erros nas variáveis.

A escolha desse conjunto de dados foi feita de forma empírica, após a avaliação de diferentes bases, com o objetivo de selecionar uma que evidenciasse de forma clara a diferença entre as decisões obtidas com a estatística da razão de verossimilhanças tradicional (LR) e suas versões corrigidas por Bartlett. Essa abordagem busca destacar, de maneira prática, a relevância da correção proposta, especialmente em situações com tamanho amostral reduzido.

Com base em outras análises, estima-se que a variância do erro para a contagem da primavera seja de, aproximadamente, seis vezes maior do que em agosto. Assim, consideramos $\lambda_e = 6^{-1}$ conhecida. Logo, o conjunto de parâmetros a ser estidado é $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)^\top = (\beta, \alpha, \mu_x, \sigma_x^2, \sigma_u^2)^\top$. Após a utilização do método de máxima verossimilhança, obtemos $\hat{\beta} = 0,752$, $\hat{\alpha} = 1,116$, $\hat{\mu}_x = 10,047$, $\hat{\sigma}_x^2 = 2,920$ e $\hat{\sigma}_u^2 = 0,459$.

Agora nosso objetivo é verificar se β pode ser 1, para isso vamos proceder o teste de hipóteses dado por $\mathcal{H}_0 : \beta = 1$. As estatísticas de teste utilizadas são: a razão de verossimilhanças (LR) e suas versões corrigidas por *Bartlett* (LR_{B_1}, LR_{B_2} e LR_{B_3}). O valor da estatística, juntamente com o valor-p, são apresentados na Tabela 2.

Tabela 2 – Resumo das estatística do teste $\mathcal{H}_0$: $\beta = 1$

Estatística	**Valor**	**Valor-p**
LR	4,172	0,041
LR_{B_1}	2,320	0,128
LR_{B_2}	1,880	0,170
LR_{B_3}	0,845	0,358

Fonte: Autores (2023)

Observamos que a estatística da razão de verossimilhanças tradicional (LR) apresenta valor-p$= 0,041$, indicando assim a rejeição de $\mathcal{H}_0 : \beta = 1$, ao nível de 5% de significância. Enquanto que as estatísticas corrigidas indicam a não rejeição da hipótese nula, também ao nível de significância de 5%. De acordo com toda a teória apresentada e o estudo de simulação destacado, tal resultado não é surpreendente,

visto que para tamanho amostral pequeno, a estatística tradicional apresenta maior proporção de rejeição da hipótese nula do que o valor do nível de significância especificado.

Os resultados apresentados têm implicações práticas relevantes, especialmente em contextos nos quais a amostra disponível é pequena e os dados estão sujeitos a erros de mensuração. Por exemplo, em estudos médicos, nos quais coletas de dados com alta precisão podem ser difíceis e onerosas, o uso dos testes corrigidos pela metodologia proposta pode proporcionar inferências mais confiáveis. O mesmo se aplica a áreas como a química analítica ou as ciências ambientais, nas quais variáveis explicativas frequentemente apresentam incertezas. Dessa forma, a adoção da correção de Bartlett pode melhorar substancialmente a qualidade das decisões baseadas em testes estatísticos, contribuindo para análises mais robustas em diversas áreas aplicadas.

6 ROTINA PARA O CÁLCULO DA ESTATÍSTICA CORRIGIDA

Com o objetivo de facilitar a utilização do modelo com erros nas variáveis, juntamente com o teste da razão de verossimilhanças corrigido, construímos uma rotina na linguagem R Development Core Team (2022) e disponibilizamos no repositório GitHub `https://github.com/Kaique-fhs/bartlett`. A rotina apresenta apenas a função `MEM`, mas com ela é possível obter as estimativas dos parâmetros do modelo com erros nas variáveis, juntamente com a estatística da razão de verossimilhanças (`LR`) e suas versões corrigidas por *Bartlett* (`LR_B1`, `LR_B2` e `LR_B3`). Os argumentos da função são:

- `formula`: um objeto da classe “formula”, sendo uma descrição simbólica do modelo a ser ajustado.

- `lambda_x`: um escalar, sendo dado pela razão entre Var(x_i) e Var(u_i), com *default* igual a *FALSE*.

- `lambda_e`: um escalar, sendo dado pela razão entre Var(e_i) e Var(u_i), com *default* igual a *FALSE*. É necessário que `lambda_x` ou `lambda_e` seja declarado.

- `beta_til`: um escalar, valor da hipótese nula do teste da razão de verossimilhanças, ou seja $\mathcal{H}_0 : \beta =$ `beta_til`.

- `Correction`: um valor lógico, se `TRUE`, a saída irá imprimir as estatísticas corrigidas e os respectivos valor-p. Mas se `FALSE`, as estatísticas corrigidas não serão apresentadas.

Uma alternativa para acessar as rotinas no repositório GitHub é utilizando o pacote `devtools`. A exemplificação da rotina usando os dados descritos na Seção 5 é apresentada a seguir:

```
> install.packages("devtools")
> library(devtools)
> install_github("Kaique-fhs/bartlett")
> library(bartlett)
> x <- c(9, 6.6, 12.3, 11.9, 11.9, 12, 9.6, 7.5, 10.9, 10.4, 10.2, 7.4,
11, 11.8, 8.2)
> y <- c(8, 6, 9.8, 10.8, 9.7, 9.3, 9.2, 6.9, 8.1, 8.7, 8.7, 7.4, 10.1,
10, 7.3)
> MEM(y ~ x,  lambda_e = 1/6,  Correction = TRUE)
Call:

MEM( formula = y ~ x,  = lambda_e = 0.1666667  Correction = TRUE)

Coefficients:
            .  Estimate  Estimate_H0      LR  P_Value_LR
 (Intercept)    1.1158        -1.38  4.1718    4.11e-02
           x    0.7516         1.00     ---           *

Bartlett's correction:
 Statistics   Value   P_Value  .
         LR  4.1718  4.11e-02  *
      LR_B1  2.3209  1.28e-01
      LR_B2  1.8792   1.7e-01
      LR_B3  0.8448  3.58e-01
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Multiple R-squared: 0.826 , Mean of x: 10.04667 , Mean of y: 8.666667
Sigma_x: 2.92 considering the null hypothesis: 1.787
Sigma_u: 0.459 considering the null hypothesis: 0.614
```

Como é possível de se observar, a utilização da rotina disponibilizada é um facilitador para a utilização do teste da razão de verossimilhanças corrigida, quando o modelo com erros nas variáveis é empregado para amostras pequenas ou moderadas.

7 CONCLUSÕES

Neste trabalho obtivemos expressões matriciais para a correção de *Bartlett* no modelo de regressão linear simples com erros nas variáveis. A expressão encontrada para este ajuste é uma fórmula que envolve operações de vetores e matrizes. Comparamos os desempenhos dos testes baseados nas estatísticas da razão de verossimilhanças original e corrigida, através de simulação de Monte Carlo. Observamos que o teste corrigido proposto apresenta desempenho superior ao teste tradicional, especialmente em amostras pequenas. Usamos um conjunto de dados reais para ilustrar a teoria estudada, o que resultou em decisões distintas entre as versões original e corrigidas da estatística. Além disso, disponibilizamos uma rotina para a correção de *Bartlett* em linguagem `R`, no repositório `GitHub`, facilitando sua aplicação prática.

Os achados deste estudo reforçam a efetividade da correção de Bartlett na melhoria da qualidade inferencial em amostras pequenas e demonstram sua viabilidade prática em modelos com erros nas variáveis — uma abordagem ainda pouco explorada em aplicações empíricas. A disponibilização da rotina em `R` amplia o potencial de uso da metodologia, tornando-a acessível a pesquisadores de diferentes áreas. Além disso, os resultados estão em consonância com a literatura, como demonstrado por Cordeiro & Cribari-Neto (1996, 2014), Bayer & Cribari-Neto (2013), Noma et al. (2018) e Melo et al. (2022), que evidenciam os benefícios de correções assintóticas de ordem superior para a inferência estatística.

A principal contribuição deste trabalho está na adaptação e explicitação das expressões de correção de Bartlett para o modelo com erros nas variáveis em regressão linear simples, sob diferentes suposições de identificabilidade, além da disponibilização de uma rotina computacional que facilita a aplicação prática desse ajuste.

AGRADECIMENTOS

Agradecemos aos revisores pelos comentários e sugestões.

REFERENCES

Aoki, R., Bolfarine, H., & Singer, J. M. (2001). Null intercept measurement error regression models. *Test*, 10(A):441–457.

Araújo, M. C., Cysneiros, A. H. M. A., & Montenegro, L. C. (2022). Bartlett and bartlett-type corrections in heteroscedastic symmetric nonlinear regression models. *An Acad Bras Cienc*, 21:1–15.

Aubin, E. C. Q. & Cordeiro, G. M. (2000). Bartlett-corrected tests for normal linear models when the error covariance matrix is nonscalar. *Communications in Statistics – Theory and Methods*, 29:2405–2426.

Barndorff-Nielsen, O. (1986). Inference on full or partial parameters, based on the standardized signed log likelihood ratio. *Biometrika*, 73:307–322.

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London A*, 160:268–282.

Bartlett, M. S. (1947). Multivariate analysis. *Journal of the Royal Statistical Society. Supplement*, 9:179–197.

Bartlett, M. S. (1954). A note on the multiplying factors for various χ^2 approximations. *Journal of the Royal Statistical Society B*, 16:296–298.

Bayer, F. M. & Cribari-Neto, F. (2013). Bartlett corrections in beta regression models. *Journal of Statistical Planning and Inference*, 143:531–547.

Chan, L. K. & Mak, T. K. (1979). On the maximum likelihood estimation of a linear structural relationship when the intercept is known. *Journal of Multivariate Analysis*, 9:304–313.

Cheng, C. L. & Van Ness, J. W. (1999). *Statistical Regression with Measurement Error*. Oxford University Press.

Cordeiro, G. M. (1993). General matrix formulae for computing bartlett corrections. *Statistics and Probability Letters*, 16:11–18.

Cordeiro, G. M. & Cribari-Neto, F. (1996). On bartlett and bartlett-type corrections. *Econometric Reviews*, 15:339–367.

Cordeiro, G. M. & Cribari-Neto, F. (2014). *An Introduction to Bartlett Correction and Bias Reduction*. Springer.

Fuller, W. A. (1987). *Measurement Error Models*. Wiley.

Kulathinal, S. B., Kuulasmaa, K., & Gasbarra, D. (2002). Estimation of an errors-in-variables regression model when the variances of the measurement errors vary between the observations. *Statistics in Medicine*, 21:1089–1101.

Magalhães, T. M., Pereira, G. H. A., Botter, D. A., & Sandoval, M. C. (2024). Bartlett corrections for zero-adjusted generalized linear models. *Statistical Papers*, 65:2191–2209.

Melo, T. F. N., Vargas, T. M., Lemonte, A. J., & Moreno-Arenas, G. (2022). Higher-order asymptotic refinements in the multivariate dirichlet regression model. *Communications in Statistics – Simulation and Computation*, 51:53–71.

Melo, T. F. N., Vargas, T. M., Lemonte, A. J., & Patriota, A. G. (2024). Higher-order asymptotic refinements in a multivariate regression model with general parameterization. *Journal of Statistical Computation and Simulation*, 94(13):2952–2975.

Neyman, J. & Pearson, E. S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika*, 20:175–240.

Noma, H., Nagashima, K., Maruo, K., Gosho, M., & Furukawa, T. A. (2018). Bartlett-type corrections and bootstrap adjustments of likelihood-based inference methods for network meta-analysis. *Statistics in Medicine*, 37:1178–1190.

R Development Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Riu, J. & Rius, F. X. (1996). Assessing the accuracy of analytical methods using linear regression with errors in both axes. *Analytical Chemistry*, 68:1851–1857.

Wong, M. Y. (1991). Bartlett adjustment to the likelihood ratio statistic for testing several slopes. *Biometrika*, 78:221–224.

Wong, O., Qiao, G., Di Loreto, A., Shi, X., & Wong, A. (2025). Bartlett-type correction for testing homogeneity of inverse gaussian means. *Journal of Statistical Computation and Simulation*, pages 1–20.

Zavala, A. A. Z., Bolfarine, H., & De Castro, M. (2007). Assessing the accuracy of analytical methods using linear regression with errors in both axes. *Annals of the Institute of Statistical Mathematics*, 59:515–530.

Contribuições dos autores

1 – Kaíque Ferreira Henrique de Souza
Bacharel em Estatística pela Universidade Federal de Goiás
https://orcid.org/0009-0006-0126-5635 • kaiquefhds@gmail.com
Contribuição: Curadoria dos dados, Investigação, Metodologia, Software, Validação, Visualização e Redação – preparação do rascunho original

2 – Tatiane Ferreira do Nascimento Melo (Corresponding Author)
Doutora em Estatística pela Universidade de São Paulo
https://orcid.org/0000-0001-5035-2280 • tmelo@ufg.br
Contribuição: Conceituação, Curadoria dos dados, Análise formal, Investigação, Metodologia, Administração do projeto, Software, Supervisão, Validação, Visualização, Redação – preparação do rascunho original e Redação – revisão e edição

3 – Eder Angelo Milani
Doutor em Estatística pela Universidade de São Carlos
https://orcid.org/0000-0001-5533-6693 • edermilani@ufscar.br
Contribuição: Análise formal, Investigação, Metodologia, Software, Supervisão, Validação, Visualização, Redação – preparação do rascunho original e Redação – revisão e edição

4 – Tiago Moreira Vargas
Doutor em Estatística pela Universidade de São Paulo
https://orcid.org/0000-0003-3697-5101 • vargas@ufg.br
Contribuição: Conceituação, Curadoria dos dados, Análise formal, Investigação, Metodologia, Administração do projeto, Software, Supervisão, Validação, Visualização, Redação – preparação do rascunho original e Redação – revisão e edição

Como citar este artigo

Souza, K. F. H., Melo, T. F. N., Milani, E. A., & Vargas, T. M. (2026). Bartlett's correction in measurement errors models. *Ciência e Natura*, Santa Maria, v. 48, e83861. DOI: https://doi.org/10.5902/2179460X83861.