Estatísticas da Razão de Verossimilhança, Wald e Escore em amostras pequenas para a distribuição Beta.

Wesley Oliveira Furriel[1], André Felipe B. Menezes[2] e [3]Rosangela Getirana Santana

[1,2,3] Universidade Estadual de Maringá

RESUMO

Sabendo-se a importância dos indicadores sociais e da saúde para compreensão dos fenômenos que cercam a sociedade contemporânea e a flexibilidade da distribuição de probabilidade Beta para sua modelagem. Este trabalho investigou as estatísticas da Razão de Verossimilhança, Wald e Escore, analisando a probabilidade do *Erro do Tipo I* (α), com base na condução de um estudo de simulação Monte Carlo, visando avaliá-los em distintos cenários, variando o tamanho das amostras (n) e os valores dos parâmetros μ e ϕ. Para a execução desse procedimento adotou-se uma reparametrização da Beta proposta por Ferrari e Cribari-Neto (2004) presente na análise de regressão. Os resultados obtidos revelaram diferenças entre os testes em relação ao α, quando o tamanho da amostra é pequeno, bem como melhores resultados no panorama geral, para o teste da Razão de Verossimilhança.
Palavras chave: Distribuição Beta, simulação Monte Carlo, Erro do Tipo I, teste da Razão de Verossimilhança, teste de Wald, teste Escore.

1 Introdução

Nas mais variadas áreas do conhecimento o emprego de indicadores, proporções ou taxas tem sido amplamente utilizado para a sumarização, identificação e hierarquização de distintos fenômenos analisados pelos pesquisadores. Isso ocorre devido ao fato de que, variáveis que assumem um intervalo definido entre um valor mínimo e um máximo, permitem diagnósticos rápidos a cerca de determinado evento, uma vez que, elas podem, por exemplo, indicar o quão adequado ou inadequado determinado caso está na medida em que ele se aproxima dos valores presentes em umas destas extremidades. Exemplos desses casos são o IDH (Índice de Desenvolvimento Humano) e a Cobertura Vacinal.

Para a modelagem de casos como os descritos acima, a família de distribuições Beta é apresentada como uma das alternativas mais adequadas.

No que diz respeito a esta distribuição, Ferrari e Cribari-Neto (2004) sugeriram uma reparametrização desta distribuição para sua utilização em modelos de regressão. Já que, na análise de regressão, geralmente é mais útil modelar resposta média, assim, ele segue as propriedades da Beta, sendo adequado para casos em que a variável resposta Y é medida continuamente no intervalo $0 < Y < 1$. Seguindo essa parametrização, $\mu = \alpha/(\alpha + \beta)$ e $\phi = \alpha + \beta$, ou seja, $\alpha = \mu\,\phi$ e $\beta = (1 - \mu)\,\phi$. Isto posto, a média e a variância de Y são dadas pelas seguintes expressões,

$$\mathbb{E}(Y) = \mu \quad \text{e} \quad \mathrm{Var}(Y) = \frac{\mathrm{V}(\mu)}{(1 + \phi)} \tag{1}$$

em que $\mathrm{V}(\mu) = \mu(1 - \mu)$, de modo que μ é a média da variável resposta e ϕ pode ser interpretado como um parâmetro de dispersão. Tendo em vista essa nova parametrização em termos de μ e ϕ, Y tem-se uma função de densidade Beta dada por:

$$f(y \mid \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi - 1}(1 - y)^{(1-\mu)\phi - 1}, \; 0 < y < 1 \tag{2}$$

com $0 < \mu < 1$ e $\phi > 0$. Assim como na parametrização original é possível obter diferentes formas no comportamento da distribuição de acordo com os valores de seus parâmetros μ e ϕ. Ferrari e Cribari-Neto(2004) apontam que a distribuição pode ser simétrica quando $\mu = 1/2$ e assimétrica quando $\mu \neq 1/2$. Além disso, a dispersão da distribuição, para um μ fixado diminui quando os valores de ϕ aumentam.

Tendo em vista o panorama apresentado o presente trabalho analisou os resultados dos testes de hipótese de Razão de Verossimilhança, Wald e Escore, a partir da condução de um estudo de simulação Monte Carlo, que avaliou os testes, em distintos cenários, nos quais variou-se o tamanho das amostras (n) e os valores dos parâmetros μ e ϕ.

Os três testes citados acima são fundamentados na teoria da verossimilhança, especificamente eles avaliam a função log-verossimilhança em diferentes escalas. A estatística da Razão de Verossimilhança compara a altura das log-verossimilhança dos modelos completo e restrito, ao passo que a estatística de Wald compara a estimativa do parâmetro ($\widehat{\boldsymbol{\theta}}$) com o valor do parâmetro sob a hipótese nula ($\boldsymbol{\theta}_0$). Enquanto que, a estatística Escore verifica a inclinação da log-verossimilhança sob a hipótese nula.

Assim sendo, considere $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ uma amostra aleatória de n observações oriundas de (2), tem-se que o logaritmo natural da função verossimilhança pode ser escrito na forma:

$$\begin{aligned} l(\mu, \phi \mid \boldsymbol{y}) = & \, n \, \log \Gamma(\phi) - n \, \log \Gamma(\mu\phi) - n \, \log \Gamma(\phi - \mu\,\phi) \\ & + (\mu\,\phi - 1) \sum_{i=1}^{n} \log \, y_i + [\phi - \mu\,\phi - 1] \sum_{i=1}^{n} \log\,(1 - y_i) \end{aligned} \tag{3}$$

Diferenciando (3) em relação μ e ϕ temos as componentes do vetor escore:

$$\begin{aligned} \mathcal{U}_\mu(\boldsymbol{\theta} \mid \boldsymbol{y}) &= n\,\phi\,\Psi\,(\phi - \mu\,\phi) - n\,\phi\,\Psi\,(\mu\,\phi) + \phi \sum_{i=1}^{n} \log\, y_i - \phi \sum_{i=1}^{n} \log\,(1 - y_i)\,, \\ \mathcal{U}_\phi(\boldsymbol{\theta} \mid \boldsymbol{y}) &= n\,\Psi\,(\phi) + [n\,(\mu + \phi)]\;\Psi(\phi - \mu\,\phi) - n\,\mu\,\Psi(\mu\,\phi) + \mu \sum_{i=1}^{n} \log\, y_i + (1 - \mu) \sum_{i=1}^{n} \log\,(1 - y_i). \end{aligned}$$

em que Ψ é a função digama. Os estimadores de máxima verossimilhança de $\boldsymbol{\theta} = (\mu, \phi)$, denotados por $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}, \widehat{\phi})$, podem ser obtidos igualando o vetor escore a zero e resolvendo simultaneamente o sistema de equações não lineares.

Para a melhor sistematização e organização da discussão e dos resultados, o trabalho foi dividido em quatro seções. Na seção 2 foram apresentados os testes de hipótese que foram avaliados, na seção 3 nos ocupamos em detalhar os cenários do estudo de simulação e por fim na seção 4, foram apresentados os resultados. É preciso ressaltar que este estudo tem como intuito iniciar a discussão acerca do uso da distribuição Beta para a modelagem de indicadores sociais e da saúde, atentando-se a sua importância para a compreensão da realidade social e consequentemente para o auxílio das políticas públicas, uma vez que, pretende-se em estudos posteriores realizar uma análise de regressão tendo como variável de interesse o IDH, bem como, indicadores básicos de saúde que podem ser descritos a partir da Beta.

2 Teste de hipóteses

Nesta seção temos teoria dos testes avaliados no estudo de simulação. Seja $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ uma amostra aleatória de n observações independentes proveniente de (2), suponha o interesse em testar a hipótese nula $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ versus a alternativa $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, em que $\boldsymbol{\theta} = (\mu, \phi)$.

Através do teste da razão de verossimilhança deve-se avaliar o máximo das funções log-verossimilhanças dos modelos restrito e completo. Sob a hipótese nula a estatística:

$$T_1 = 2\left[l(\widehat{\boldsymbol{\theta}} \mid \boldsymbol{y}) - l(\boldsymbol{\theta}_0 \mid \boldsymbol{y})\right] \tag{4}$$

possui distribuição assintótica qui-quadrado com $p = 2$ graus de liberdade (PAWITAN, 2001).

Por outro lado, o teste de Wald consiste em comparar a estimativa do parâmetro $(\widehat{\boldsymbol{\theta}})$ com o valor do parâmetro sob a hipótese nula $(\boldsymbol{\theta}_0)$, por isso requer somente o ajuste do modelo completo. Como temos duas restrições, sob a hipótese nula a estatística do teste tem forma quadrática definida por (MILLAR, 2011):

$$T_2 = \left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)^T \mathcal{I}(\widehat{\boldsymbol{\theta}}) \left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \tag{5}$$

sendo $\mathcal{I}(\widehat{\boldsymbol{\theta}}) = \mathrm{Var}(\widehat{\boldsymbol{\theta}})$ a matriz de informação esperada localmente nas estimativas de máxima verossimilhança $\widehat{\mu}$ e $\widehat{\phi}$. Desse modo, sob a hipótese nula T_2 tem assintoticamente distribuição qui-quadrado com $p = 2$ graus de liberdade.

Por fim, o teste Escore baseia-se na conjectura da função escore $\mathcal{U}(\boldsymbol{\theta}_0)$ e da matriz de informação esperada $\mathcal{I}(\widehat{\boldsymbol{\theta}}_0)$ sob a hipótese nula. Conforme Millar (2011), a estatística do teste pode ser determinada em termos de forma quadrática por:

$$T_3 = \left(\mathcal{U}(\boldsymbol{\theta}_0)\right)^T \mathcal{I}(\boldsymbol{\theta}_0)^{-1}\, \mathcal{U}(\boldsymbol{\theta}_0) \tag{6}$$

em que, $\mathcal{U}(\boldsymbol{\theta}_0) = \dfrac{\partial}{\partial \boldsymbol{\theta}} l(\boldsymbol{\theta} \mid \boldsymbol{y})\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$ e $\mathcal{I}(\boldsymbol{\theta}_0)^{-1}$ é a inversa da matriz de informação esperada sob a hipótese nula. A estatística T_3, sob a hipótese nula tem assintoticamente distribuição qui-quadrado com $p = 2$ graus de liberdade.

Os autores Mazuchelli e Louzada (2014) ressaltam que estes três testes são assintoticamente equivalentes, todavia podem diferir em amostras pequenas.

3 Estudo de simulação

Diante dos testes descritos na seção anterior, que avaliam as hipóteses por meio de escalas distintas é de grande interesse verificar seu comportamento em distintas situações. Desse modo, foi conduzido um estudo de simulação Monte Carlo que averiguou o comportamento do tamanho do teste (*Erro do Tipo I*) em diferentes cenários. Assim, foram tomadas amostras de tamanhos $n = 20, 40, 60, 80$ e 100, $\mu = 0.05, 0.25, 0.5, 0.75$ e 0.95 e $\phi = 5, 15$ e 50. Para cada combinação (n, μ, ϕ) foram geradas $B = 10000$ amostras pseudoaleatórias da nova parametrização da distribuição Beta (2). Em todas as instâncias adotamos níveis de significância nominais $\alpha = 0,05$.

Para atingir os objetivos propostos foi implementado no *software* SAS uma *macro* para calcular as probabilidades de rejeitar H_0 quando a mesma foi verdadeira (valores-p) dos testes da Seção 2. Sendo que, o teste da Razão de Verossimilhança e Wald foram obtidos via `PROC NLMIXED`, um procedimento do SAS que utiliza métodos baseado em verossimilhança para o estudo de modelos não lineares e mistos. Já o teste Escore foi implementado no SAS/IML.

4 Resultados

No que tange os resultados, foi possível verificar que o teste da Razão de Verossimilhança, de um modo geral, foi o que apresentou o melhor desempenho quando comparado aos demais, já que na maior parte dos cenários sua convergência para o nível de significância nominal de $0,05$ se dá de forma mais acentuada. Além disso, verificou-se, para todos os testes que quando o tamanho da amostra é pequeno a probabilidade do *Erro Tipo I* excede o nível de significância nominal. Por fim, apontamos que para o cenário, no qual, o parâmetro de dispersão é $\phi = 5$, não obtivemos a convergência para o nível de significância nominal exigido. Desse modo, pretendemos investigar tal fenômeno em trabalhos posteriores

Referências

[1] FERRARI, S.; CRIBARI-NETO, F. Beta regression for modelling rates and proportions. Journal of Applied Statistics, v. 31, n. 7, p. 799-815, 2004.

[2] MAZUCHELI, J.; LOUZADA, F. Discriminação entre as distribuições Odd Weibull e Weibull. Rev. Bras. Biom 32.2 (2014): 226-237.

[3] MILLAR, R. B. Maximum likelihood estimation and inference: with examples in R, SAS and ADMB. Vol. 111. John Wiley & Sons, 2011.

[4] PAWITAN, Y. In all likelihood: statistical modelling and inference using likelihood. Oxford University Press, 2001.