Revista da Estatística UFOP, Vol III(3), 2014, ISSN 2237-8111
Edição especial: 59ª Reunião Anual da Regional Brasileira da Sociedade Internacional de Biometria - RBRAS
Trabalhos apresentados na versão poster.

Dimensionamento de amostras para estudos com variável resposta ordinal

Arminda Lucia Siqueira[1]
Aloísio Joaquim Freitas Ribeiro[1]
Edna Afonso Reis[1]
Ilka Afonso Reis[1]

1 Introdução

Uma situação comum na prática é quando há o interesse em avaliar a relação entre uma variável resposta ordinal e determinadas variáveis explicativas. Há diferentes métodos de análise de dados ordinais, incluindo os seguintes *modelos de regressão logística ordinal*: *modelo de chances proporcionais, modelo de razão-contínua, modelo estereótipo e modelo de chances proporcionais parciais.* Detalhes sobre os modelos são encontrados em Hosmer & Lemeshow (2000), Abreu *et al.* (2008), Abreu *et al.* (2009) e Hilbe (2009).

Assim como no caso de regressão logística para dados de resposta binária, em regressão logística ordinal sugere-se o uso da medida resumo conhecida por *odds ratio* (*OR*) para avaliação do efeito da variável explicativa na resposta. Suponha a resposta de interesse Y tenha k categorias ordenadas ($Y_j,\ j = 1, 2, \cdots, k$) e que se deseja comparar dois grupos (A e B). Para a categoria j, a medida OR é definida por: $OR_j = \frac{Pr\,(Y \leq Y_j\,|x_{(A)})}{1 - Pr\,(Y \leq Y_j\,|x_{(A)})} \div \frac{Pr\,(Y \leq Y_j\,|x_{(B)})}{1 - Pr\,(Y \leq Y_j\,|x_{(B)})}$. Em palavras, OR é a razão entre a probabilidade cumulativa de estar na categoria j ou inferior dentro do grupo A e o complementar dessa probabilidade, dividida pela razão das mesmas probabilidades cumulativas referentes ao grupo B.

Uma questão relevante é sobre o dimensionamento da amostra para estudos cuja variável resposta é ordinal. Neste trabalho, apresentamos uma revisão da literatura sobre o assunto destacando o modelo de chances proporcionais e um método não paramétrico. Ilustramos a metodologia com exemplos de dados simulados e de situações reais.

[1] Departamento de Estatística/ UFMG. e-mail: arminda@est.ufmg.br

2 Material e métodos

Os dados consistem no vetor de variáveis explicativas $\boldsymbol{x} = (x_1, x_2, \cdots x_p)$ e na variável resposta Y com k categorias ordenadas ($Y_j,\ j = 1, 2, \cdots, k$) que, condicionalmente aos valores de $\boldsymbol{x}$, ocorrem com probabilidades $p_1, p_2, \cdots, p_k$, isto é, $p_j = Pr(Y = j)$, para $j = 1, 2, \cdots, k$.

O Modelo de Chances Proporcionais (MCP), em inglês *proportional odds model*, é escrito como $log\frac{Pr(Y=1|\boldsymbol{x})+\cdots+Pr(Y=j|\boldsymbol{x})}{Pr(Y=j+1|\boldsymbol{x})+\cdots+Pr(Y=k|\boldsymbol{x})} = \alpha_j + \beta_1 x_1 + \cdots + \beta_p x_p,\ j = 1, \cdots, k-1$. O modelo MCP compara a probabilidade de uma resposta igual ou inferior a determinada categoria ($j = 1, \cdots, k-1$), com a probabilidade de resposta superior a esta categoria. O modelo é composto por $(k-1)$ equações lineares paralelas. No caso particular de apenas duas categorias $(k = 2)$, o modelo MCP corresponde exatamente ao tradicional modelo de regressão logística binária. Detalhes adicionais sobre o modelo MCP podem ser encontrados em Hosmer & Lemeshow (2000), Abreu *et al.* (2008), Abreu *et al.* (2009) e Hilbe (2009).

Para o modelo MCP no caso da comparação de dois grupos, (por exemplo, experimental e controle, abreviados por E e C, respectivamente), Whitehead (1993) propõe um método para dimensionar o tamanho da amostra, que tem a seguinte fórmula fechada: $n_{ordinal} = \frac{3(A+1)^2(z_{1-\alpha/2}+z_{1-\beta})^2}{A[\log(OR)]^2(1-\sum_{i=1}^{k}\bar{p}_i{}^3)}$. Na fórmula, A é a razão de alocação entre os dois grupos, α e $1-\beta$ são respectivamente o nível de significância e o poder do teste, $z_{1-\alpha/2}$ e $z_{1-\beta}$ são os percentis de ordem 1-α/2 e $1-\beta$ da distribuição normal padrão; p_j é a proporção média dos dois grupos (E e C) para a j-ésima categoria, isto é, $\bar{p}_j = (p_{Ej} + p_{Cj})/2$. OR é definida em termos das probabilidades acumuladas e, portanto, só pode ser calculado até a penúltima categoria ($j = 1,2, \cdots, k-1$) já que até a última categoria, as probabilidades acumuladas são iguais a um e, portanto, OR não é definido. Além disso, para o modelo MCP, teoricamente OR é igual para todas as categorias. Deve-se mencionar que a fórmula anterior fornece o tamanho total da amostra, não o tamanho por grupo, como em geral aparece em várias fórmulas. Frequentemente, $A = 1$, isto é, a taxa de alocação dos grupos é a mesma. Neste caso, a fórmula anterior torna-se $n_{ordinal\ (A=1)} = \frac{12(z_{1-\alpha/2}+z_{1-\beta})^2}{[\log(OR)]^2(1-\sum_{i=1}^{k}\bar{p}_i{}^3)}$ e para a fórmula para cada grupo seria $n_{ordinal\ (A=1)} = \frac{6(z_{1-\alpha/2}+z_{1-\beta})^2}{[\log(OR)]^2(1-\sum_{i=1}^{k}\bar{p}_i{}^3)}$.

Raharhja *et al.* (2009) apresenta uma revisão bastante completa sobre determinação de amostra utilizando o teste Wilcoxon-Mann-Whitney (WMW), também conhecido como teste de Mann Whitney, o teste não paramétrico mais popular para comparar dois grupos (E e C) de observações contínuas ou uma variável categórica com ordenação. Para as k categorias com ordenação $(C_1, C_2, \cdots, C_k)$, sejam m_j e n_j as frequências para a j-ésima categoria, respectivamente para os grupos E e C, com totais m e n, $N = m + n,$ e $t = n/N$ a fração entre tratamentos. As correspondentes proporções em cada categoria são $p_j = m_j /m$ e $q_j = n_j /n$, No artigo, destaca-se o trabalho de Zhao *et al.* (2008) e a seguinte fórmula para o tamanho de amostra com enfoque não paramétrico quando o pressuposto de chances proporcionais não é válido é fornecida: $n_{NP} = \frac{(z_{1-\alpha/2}+z_{1-\beta})^2(1-\sum_{c=1}^{k}((1-t)p_c+tq_c)^3)}{12t(1-t)(\sum_{c=2}^{k} p_c \sum_{d=1}^{c-1} q_d + 0{,}5\sum_{c=1}^{k} p_c q_c - 0{,}5)^2}$. Para aplicar a fórmula, é necessário conhecer t e as probabilidades (p e q) para todas as categorias. Note que esta fórmula permite frequências nulas em alguma(s) categoria(s).

3 Resultados e discussão

A seguir apresentamos exemplos simulados ou extraídos da literatura, procurando ilustrar o que foi apresentado na seção anterior. Nos primeiros exemplos, a fórmula de Whitehead (1993) e o método não paramétrico foram empregados, respectivamente. No terceiro exemplo, quatro métodos para o cálculo do tamanho de amostra e do poder foram comparados.

3.1 Dimensionamento de ensaio clínico com dois grupos

Em um ensaio clínico, as categorias da variável resposta foram classificadas como muito boa, boa, moderada e ruim; o grupo experimental foi comparado ao grupo controle. Os três conjuntos de dados são mostrados no Quadro 1 e a Tabela 1 mostra o total da amostra calculado pelo método de Whitehead (1993), para quatro taxas de alocação: $A = 1, 2, 3{,}4$.

Quadro 1: Conjuntos de Dados

Grupo	Categorias			
	Muito boa	Boa	Moderada	Ruim
(a) Dados extraídos de Whitehead (1993)				
Controle	$p_{C1} = 0{,}20$	$p_{C2} = 0{,}50$	$p_{C3} = 0{,}20$	$p_{C1} = 0{,}10$
Experimental	$p_{E1} = 0{,}378$	$p_{E2} = 0{,}472$	$p_{E3} = 0{,}106$	$p_{E4} = 0{,}044$
(b) Dados simulados – as probabilidade do grupo experimental são bem superiores				
Controle	$p_{C1} = 0{,}20$	$p_{C2} = 0{,}30$	$p_{C3} = 0{,}40$	$p_{C1} = 0{,}10$
Experimental	$p_{E1} = 0{,}55$	$p_{E2} = 0{,}15$	$p_{E3} = 0{,}25$	$p_{E4} = 0{,}05$
(c) Dados simulados – as probabilidades das categorias do grupo controle são 0,25				
Controle	$p_{C1} = 0{,}25$	$p_{C2} = 0{,}25$	$p_{C3} = 0{,}25$	$p_{C4} = 0{,}25$
Experimental	$p_{E1} = 0{,}30$	$p_{E2} = 0{,}28$	$p_{E3} = 0{,}27$	$p_{E4} = 0{,}15$

Tabela 1: Total da amostra segundo taxa de alocação

Conjunto de dados	Taxa de alocação			
	$A = 1$	$A = 2$	$A = 3$	$A = 4$
(a)	187	211	250	292
(b)	195	200	237	278
(c)	1.295	1.457	1.726	2.023

À medida que a taxa de alocação aumenta, o tamanho total da amostra também cresce, mas não de forma proporcional. Por exemplo, para o conjunto de dados (a), se $A = 2$, $n_{ordinal}$ é apenas $1{,}13$ o tamanho correspondente ao de $A = 1$ ($211/187$=1,13). Whitehead (1993) apresenta uma seção sobre considerações de planejamento em que avalia a taxa de alocação A. De forma geral, concluiu que há uma relação entre o tamanho de amostra e o número de categorias e que a taxa de alocação excedendo quatro raramente se justifica.

3.2 Dimensionamento de amostra por método não paramétrico (Zhao *et al.*, 2008)

Para os dados sobre a associação entre fumo e condição da retinopatia em pacientes diabéticos, apresentados na Tabela 2, o pressuposto do modelo de chances proporcionais não é validado ($p = 0{,}017$ para o teste de ajuste do modelo).

Quando $t = 325/613 = 0{,}53$, mantendo a mesma distribuição das categorias do grupo dos não fumantes (66%, 15%, 19%), mas trocando os percentuais das categorias do grupo dos fumantes para (55%, 20%, 25%) ou para (55%, 15%, 30%), os valores obtidos pela fórmula do método não paramétrico para o nível de significância de 5% e poder de 80% são $n_{NP} = 671$ e $n_{NP} = 502$, respectivamente.

Tabela 2: Distribuição de frequências da condição da retinopatia segundo hábito tabagista

Fumante	Condição da retinopatia			Total
	Nenhuma	Não proliferativa	Avançada	
Não	191 (66%)	42 (15%)	55 (19%)	288 (100%)
Sim	197 (61%)	76 (23%)	52 (16%)	325 (100%)
Total	388	118	107	613

3.3 Tamanho da amostra e poder - comparação de quatro métodos (Walters, 2004)

Quatro métodos para estimar tamanho de amostra e poder foram comparados e ilustrados com estudos de saúde relacionados à qualidade de vida (ESQV) avaliados pela escala SF-36 para a comparação de dois grupos: controle e intervenção. O método 1 compara médias quando a variável segue a distribuição normal; o método 2 é o não paramétrico de Mann Whitney; o método 3 é o de Whitehead (1993) e o método 4 utiliza *bootstrap*.

São estabelecidas as condições em que cada um dos métodos seria mais indicado e as conclusões são organizadas em um fluxograma. Um dos resultados foi que, se os resultados do ESQV têm número limitado de valores discretos (inferior a sete) e/ou o percentual esperado de casos nas fronteiras é alta (0 ou 100), o método 3 de Whitehead (1993) é recomendado.

4 Conclusões e considerações finais

O modelo de chances proporcionais é indicado quando a variável resposta era originalmente uma variável contínua que, posteriormente foi agrupada (Abreu *et al*., 2008). O método de Whitehead (1993) tem sido muito empregado no dimensionamento de amostra quando se utiliza o modelo MCP. Se o pressuposto do modelo não for válido, uma opção é utilizar um método não paramétrico.

Existem excelentes *softwares* específicos para cálculo de tamanho de amostra para várias situações, como o nQuery Advisor®, Pass® e Siz®, mas até o momento não foi feita implementação para o modelo MCP. Alternativa é fazer a própria programação, por exemplo, usando a planilha Excel® ou o *software* livre R.

5 Referências

[1] ABREU M. N. S.; SIQUEIRA, A. L.; CARDOSO, C. S.; CAIAFFA, W. T. Ordinal logistic regression models: applications in quality of life studies. **Cadernos de Saúde Pública**. FIOCRUZ. v.24 Sup, p. S581 - S591, 2008.

[2] ABREU M. N. S.; SIQUEIRA, A. L.; CAIAFFA, W. T. Regressão logística ordinal em estudos epidemiológicos. **Revista de Saúde Pública.** Universidade de São Paulo. v. 43, n. 1, p. 183 - 194, 2009.

[3] HILBE, J. M. **Logistic Regression Models.** Boca Raton: Chapman & Hall/CRC. 2009. 637 p.

[4] HOSMER, W. D.; LEMESHOW, S. **Applied Logistic Regression**, New York: JohnWiley. 2.nd ed. 2000. 373 p.

[5] RAHARDJA, D.; ZHAO, Y. D.; QU. Y. Sample size determination for the Wilcoxon-Mann-Whitney test: a comprehensive review. **Statistics in Biopharmaceutical Research**, American Statistical Association. v. 1, n. 3, p. 317-322, 2009.

[6] WALTERS, S. J. Sample size and power estimation for studies with health related quality of life outcomes: a comparison of four methods using the SF-36. **Health and Quality of Life Outcomes**, 2, 26, p. 1-17, 2004.

[7] WHITEHEAD, J. Sample size calculations for ordered categorical data. **Statistics in Medicine**. John Wiley. v. 12, p. 2257-2271, 1993

[8] ZAO, Y. D.; RAHARDJA, D.; QU. Y. Sample size calculations for the Wilcoxon-Mann-Whitney test adjusting for ties. **Statistics in Medicine**. John Wiley. v. 27, p. 462-468, 2008.

Agradecimento: à Fapemig pelo apoio ao projeto de pesquisa no qual o presente trabalho está inserido.