

Estimação de variância em dados amostrais complexos

Patrick Wöhrle Guimarães

Universidade Federal de Mato Grosso – Faculdade de Economia e Instituto de Física (UFMT/FE/IF)
78060-900, Cuiabá, MT

E-mail: probabilidade@gmail.com

Rogério de Oliveira e Sá

Universidade Federal de Mato Grosso – Faculdade de Economia.
78060-900, Cuiabá, MT

E-mail: rogeriocba2@gmail.com

RESUMO: A literatura que analisa o conjunto de métodos estatísticos aplicados em amostras complexas tem apresentado diversas metodologias para determinação da variância. Visando sintetizar essas metodologias e apresentar a sua aplicabilidade em desenhos amostrais complexos (peso, estrato e unidade primária), esse artigo estima uma série de medidas de variância associadas aos indicadores de desigualdade do Estado de Mato Grosso (renda, pobreza e riqueza) por métodos de linearização (MLT e MLR) e replicação (MRJ, MRBRR e MRB). Para tanto faz uso dos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) para o ano de 2007.

Palavras-chave: Estimação Variância, Amostra Complexa, Linearização, Replicação, PNAD.

1. INTRODUÇÃO

A literatura que analisa o conjunto de métodos estatísticos aplicados em amostras complexas (*Statistical Methods for Complex Survey Data*) tem ressaltado à necessidade de se considerar todas as informações do desenho amostral (peso, estrato e unidade primária) na realização de qualquer inferência. No processo de inferência freqüentemente há a necessidade de se mensurar a precisão das estimativas e isso é feito via determinação de uma medida de variância.

Essa determinação das medidas de variância pode ser feita levando em conta uma grande variedade de metodologias [2]. No caso específico em que a determinação dos estimadores (para os quais se deseja calcular a variância) seja não-linear há duas classes de estimadores de variância [4]: (a) de linearização em que o estimador não-linear é aproximado por um linear, com o propósito de estimar a variância (ex: Linearização de Taylor (MLT) e a Linearização de Rao (MLR)); (b) de replicação em que diversas estimativas do parâmetro populacional em estudo são calculadas a partir de diferentes partes da amostra original e depois reunidas para obter a estimativa de variância desejada (ex: a replicação de Jackknife (MRJ), a replicação por *Balanced Repeated Replication* (MRBRR) e a replicação por Bootstrap (MRB)).

Diversos estudos ([1], [3], [4]) mostraram que não há grandes diferenças ao estimar a variância pelo Método de Linearização de Taylor (MLT) e o método de replicação de Jackknife (MRJ), quando se trabalha com amostras com grande número de unidades primárias de amostragem (UPAs).

Visando investigar essa hipótese da indiferença de escolhas para determinação da medida de variância, esse artigo estima uma série de medidas de precisão dos estimadores (variância) associadas aos indicadores de desigualdade do Estado de Mato Grosso (renda, pobreza e riqueza) por métodos de linearização (MLT e MLR) e replicação (MRJ, MRBRR e MRB). Para tanto faz uso dos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) para o ano de 2007.

REFERÊNCIAS BIBLIOGRÁFICAS:

- [1] LEE, E. S.; FORTHOFFER, R. N.; LORIMOR, R. J. Analysis of complex sample survey data: problems and strategies. **Sociological Methods & Research**, v.15, n.1-2, p.69-100, 1986.
- [2] PESSOA, D. G. C.; NASCIMENTO SILVA, P. L. **Análise de dados amostrais complexos**. São Paulo: Associação Brasileira de Estatística, 1998. 170p.
- [3] RUST, K. Variance estimation for complex estimators in sample surveys. **Journal of Official Statistics**, v.1, n.4, p.381-397, 1985.
- [4] VIEIRA, M. T. Um estudo comparativo das metodologias de modelagem de dados amostrais complexos – uma aplicação ao SAEB 99. Rio de Janeiro, 2001. 108p. Dissertação (Mestrado) – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.