

Đánh giá cách các công cụ ước tính hồi quy tổng quát góp phần ước tính dữ liệu tài chính và kinh tế khi thiếu các quan sát trong lấy mẫu xác suất không đồng đều

Lawson Nuanpan

*Khoa Thống kê Ứng dụng, Đại học Công nghệ King Mongkut Bắc Bangkok,
Bangkok, Thái Lan*

Tóm tắt

Mục đích - Biết trước thông tin tài chính và kinh tế mang lại lợi ích trong việc hoạch định và xây dựng chính sách cho mọi quốc gia, đặc biệt đối với một quốc gia đang phát triển như Thái Lan và các nước châu Á khác. Thật không may, việc thiếu dữ liệu hoặc không phản hồi đóng một vai trò thiết yếu trong nhiều lĩnh vực nghiên cứu bao gồm tài chính và kinh tế. Việc loại bỏ dữ liệu còn thiếu một cách thích hợp trước khi phân tích sâu hơn có thể đạt được kết quả đáng chú ý và có thể có hiệu quả cho việc hoạch định chính sách. Đánh giá này về các ước tính hồi quy tổng quát cho tổng dân số có thể được áp dụng cho dữ liệu tài chính, kinh tế và các dữ liệu khác khi thiếu dữ liệu.

Thiết kế/phương pháp/cách tiếp cận - Các ước lượng hồi quy tổng quát để ước lượng tổng thể, bao gồm các ước lượng phương sai theo lấy mẫu xác suất không bằng nhau mà không thay thế bằng dữ liệu bị thiếu, được khám phá trong khung ngược. Các ứng dụng về dữ liệu tài chính và kinh tế ở Thái Lan cũng được xem xét.

Những phát hiện - Việc xem xét các tài liệu liên quan đến công cụ ước tính được đề xuất cho thấy hiệu suất tốt nhất, mang lại sai số nhỏ hơn trong tất cả các kịch bản.

Tính nguyên bản/giá trị - Công cụ ước tính hồi quy tổng quát có thể hỗ trợ ước tính dữ liệu tài chính và kinh tế chứa các giá trị bị thiếu với các cơ chế thiếu khác nhau và có thể được sử dụng trong các ứng dụng khác giúp thu được các công cụ ước tính ưu việt hơn.

Từ khóa Công cụ ước tính hồi quy tổng quát, Dữ liệu bị thiếu, Dữ liệu tài chính, Dữ liệu kinh tế, Lấy mẫu xác suất không đồng đều mà không thay thế

Tài liệu tham khảo

- Berger, Y.G., Tirari, E.H.M. and Till, Y. (2003), "Towards optimal regression estimation in sample surveys", *Australian and New Zealand Journal of Statistics*, Vol. 45 No. 3, pp. 319-329, doi: 10.1111/1467-842x.00286.
- Bethlehem, J.G. and Keller, W.J. (1987), "Linear weighting of sample survey data", *Journal of Official Statistics*, Vol. 3 No. 2, pp. 141-153.
- Brewer, K.R.W. (2002), *Combined Survey Sampling Inference: Weighing Basu's Elephants*, Arnold, London.
- Brewer, K.R.W. and Donadio, M.E. (2003), "The high entropy variance of the Horvitz-Thompson estimator", *Survey Methodology*, Vol. 29 No. 2, pp. 189-196.
- Deville, J.C. and Särndal, C.E. (1994), "Variance estimation for the regression imputed Horvitz Thompson estimator", *Journal of Official Statistics*, Vol. 10 No. 4, pp. 381-394.
- Estevao, V.M. and Särndal, C.E. (2003), "A new perspective on calibration estimators", *JSM- Section on Survey Research Methods*, pp. 1346-1356.
- Fay, R.E. (1991), "A design-based perspective on missing data variance", *Proceedings of the 1991 Annual Research Conference, US Bureau of the Census*, pp. 429-440.
- Hajek, J. (1964), "Asymptotic theory of rejective sampling with varying probabilities from a finite population", *Annals of Mathematical Statistics*, Vol. 35 No. 4, pp. 1491-1523, doi: 10.1214/aoms/1177700375.
- Hajek, J. (1981), *Sampling from Finite Population*, Marcel Dekker, New York.
- Hansen, M.H. and Hurwitz, W.N. (1946), "The problem of nonresponse in sample surveys", *Journal of the American Statistical Association*, Vol. 41 No. 236, pp. 517-529, doi: 10.1080/01621459.1946.10501894.
- Hartley, H.O. and Rao, J.N.K. (1962), "Sampling with unequal probability and without replacement", *The Annals of Mathematical Statistics*, Vol. 33 No. 2, pp. 350-374, doi: 10.1214/aoms/1177704564.

- Haziza, D. (2010), "Resampling methods for variance estimation in the presence of missing survey data", Proceedings of the Annual Conference of the Italian Statistical Society.
- Haziza, D. and Rao, J.N.K. (2006), "A nonresponse model approach to inference under imputation for missing survey data", Survey Methodology, Vol. 32 No. 1, pp. 53-64.
- Horvitz, D.F. and Thompson, D.J. (1952), "A generalization of sampling without replacement from a finite universe", Journal of the American Statistical Association, Vol. 47 No. 260, pp. 663-685, doi: 10.1080/01621459.1952.10483446.
- Lawson, N. (2017), "Variance estimation in the presence of nonresponse under probability proportional to size sampling", Proceedings of the 6th Annual International Conference on Computational Mathematics, Computational Geometry and Statistics (CMCGS 2017), Singapore, doi: 10.5176/ 2251-1911_cmcs17.32.
- Lawson, N. and Ponkaew, C. (2019), "New generalized regression estimator in the presence of nonresponse under unequal probability sampling", Communications in Statistics -Theory and Methods, Vol. 48 No. 10, pp. 2483-2498, doi: 10.1080/03610926.2018.1465091.
- Lawson, N. and Siripanich, P. (2022), "A new generalized regression estimator and variance estimation for unequal probability sampling without replacement for missing data", Communications in Statistics -Theory and Methods, Vol. 51 No. 18, pp. 6296-6318, doi: 10.1080/03610926.2020. 1860224.
- Midzuno, H. (1952), "On the sampling system with probability proportional to sum of sizes", Annals of the Institute of Statistical Mathematics, Vol. 55 No. 3, pp. 99-107.
- Montanari, G. (1987), "Post sampling efficient qr-prediction in large sample survey", International Statistics, Vol. 55 No. 2, pp. 191-202, doi: 10.2307/1403195.
- Ponkaew, C. and Lawson, L. (2023), "New generalized regression estimators using a ratio method and its variance estimation for unequal probability sampling

without replacement in the presence of nonresponse", *Current Applied Science and Technology*, Vol. 23 No. 2, doi: 10.55003/cast.2022.02. 23.007.

Rao, J.N.K. (1990), "Variance estimation under imputation for missing data", Technical report, Statistics Canada, Ottawa, pp. 599-608.

Särndal, C.E. (1992), "Method for estimating the precision of survey estimates when imputation has been used", *Survey Methodology*, Vol. 18, pp. 241-252.

Särndal, C.E. (2007), "The calibration approach in survey theory and practice", *Survey Methodology*, Vol. 33 No. 2, pp. 99-119.

Särndal, C.E. and Lundström, S. (2005), *Estimation in Surveys with Nonresponse*, John Wiley & Sons, New York.

Särndal, C.E., Swensson, B. and Wretman, J. (1992), *Model Assisted Survey Sampling*, Springer- Verlag, New York.

Sen, A.R. (1953), "On the estimate of the variance in sampling with varying probabilities", *Journal of the Indian Society of Agricultural Statistics*, Vol. 5, pp. 119-127.

Shao, J. and Steel, P. (1999), "Variance estimation for survey data with composite imputation and nonnegligible sampling fractions", *Journal of the American Statistical Association*, Vol. 94 No. 445, pp. 254-265, doi: 10.2307/2669700.

Yates, F. and Grundy, P.M. (1953), "Selection without replacement from within strata with probability proportional to size", *Journal of the Royal Statistical Society: Series B*, Vol. 15 No. 2, pp. 235-261, doi: 10.1111/j.2517-6161.1953.tb00140.x.

Liên hệ tác giả

Có thể liên hệ với tác giả Nuanpan Lawson tại: nuanpan.n@sci.kmutnb.ac.th