Davison, A. C.; **Sardy, S.**

Resampling methods for estimating variance in surveys. (Méthodes de rééchantillonnage pour l'estimation de variance en sondage.) (French. English summary) Zbl 1409.62032

J. Soc. Fr. Stat. 147, No. 3, 3-32 (2006).

Summary: We review resampling techniques used for variance estimation in sample surveys under calibration and imputation of missing values. The techniques considered are based on linearisation, the jackknife, balanced repeated replication, and the bootstrap. Our purpose is to give practical recommendations based on theoretical considerations and empirical comparisons. Among the resampling methods, the flexibility and stability of the bootstrap make it preferable, while for massive surveys linearisation methods can reduce the computational effort considerably.

MSC:

62D05 Sampling theory, sample surveys
62F40 Bootstrap, jackknife and other resampling methods
62-02 Research exposition (monographs, survey articles) pertaining to statistics

Keywords:

bootstrap; calibration; imputation; jackknife; linearisation; missing values; non-response; sample survey

References:

[1] BERGER, Y. G. et RAO, J. N. K. (2006) Adjusted jackknife for imputation under unequal probability sampling without replacement. Journal of the Royal Statistical Society series B 68, 531-547. · Zbl 1110.62005

[2] BERGER, Y. G. et SKINNER, C. J. (2003) Variance estimation for a low income proportion. Applied Statistics 52, 457-468. · Zbl 1111.62382

[3] BERGER, Y. G. et SKINNER, C. J. (2005) A jackknife variance estimator for unequal probability sampling. Journal of the Royal Statistical Society, Séries B 67, 79-89. · Zbl 1060.62008

[4] BICKEL, P. J. et FREEDMAN, D. A. (1984) Asymptotic normality and the bootstrap in stratified sampling. Annals of Statistics 12, 470-482. · Zbl 0542.62009

[5] BINDER, D. A. (1996) Linearization methods for single phase and two-phase samples : A cookbook approach. Survey Methodology 22, 17-22.

[6] BOONSTRA, H. J. H. et NIEUWENBROEK, N. (2003) An empirical comparison of BRR and linearization variance estimators. Centraal Bureau voor de Statistiek, The Netherlands.

[7] BOOTH, J. G., BUTLER, R. W. et HALL, P. (1994) Bootstrap methods for finite populations. Journal of the American Statistical Association 89, 1282-1289. · Zbl 0812.62006

[8] CAMPBELL, C. (1980) A different view of finite population estimation. In Proceedings of the Section on Survey Research Methods, pp. 319-324. Alexandria, Virginia : American Statistical Association.

[9] CANTY, A. J. et DAVISON, A. C. (1999) Resampling-based variance estimation for labour force surveys. The Statistician 48, 379-391.

[10] CHAO, M. T. et LO, S. H. (1985) A bootstrap method for finite populations. Sankyã A 47, 399-405. · Zbl 0587.62031

[11] CHEN, J. et SHAO, J. (2001) Jackknife variance estimation for nearest-neighbor imputation. Journal of the American Statistical Association 96, 260-269. · Zbl 1014.62008

[12] COCHRAN, W. G. (1977) Sampling Techniques. Third edition. New York : Wiley. · Zbl 0051.10707

[13] DAVISON, A. C. et HINKLEY, D. V. (1997) Bootstrap Methods and Their Application. Cambridge : Cambridge University Press. · Zbl 0886.62001

[14] DEVILLE, J. C. (1999) Variance estimation for complex statistics and estimators : Linearization and residual techniques. Survey Methodology 25, 193-203.

[15] DEVILLE, J. C. et SÄRNDAL, C. E. (1992) Calibration estimators in survey sampling. Journal of the American Statistical Association 87, 376-382. · Zbl 0760.62010

[16] DEVILLE, J. C. , SÄRNDAL, C. E. et SAUTORY, O. (1993) Generalized raking procedures in survey sampling. Journal of the American Statistical Association 88, 1013-1020. · Zbl 0794.62005

[17] DIPPO, C. S., FAY, R. E. et MORGANSTEIN, D. H. (1984) Computing variances from complex samples with replicate weights. In Proceedings of the Section on Survey Research Methods, pp. 489-494. Washington DC : American Statistical

Association.

[18] DROESBEKE, J.-J., FICHET, B. et TASSI, P. (eds) (1987) Les sondages. First edition. Economica.

[19] EFRON, B. (1979) Bootstrap methods : Another look at the jackknife. Annals of Statistics 7, 1-26. · Zbl 0406.62024

[20] EFRON, B. (1982) The Jackknife, the Bootstrap, and other Resampling Plans. Philadephia : SIAM. · Zbl 0496.62036

[21] FAY, R. E.(1989) Theory and application of replicate weighting for variance calculations. In Proceedings of the Social Statistics Section, pp. 212-217. American Statistical Association.

[22] FAY, R. E. (1996) Alternative paradigms for the analysis of imputed survey data. Journal ofthe American Statistical Association 91, 490-498. · Zbl 0869.62015

[23] GROSS, S. (1980) Median estimation in sample surveys. In Proceedings of the Section on Survey Research Methods, pp. 181-184. Alexandria, VA : American Statistical Association.

[24] GUPTA, V. K. et NIGAM, A. K. (1987) Mixed orthogonal arrays for variance estimation with unequal numbers of primary selections per stratum. Biometrika 74, 735-742. · Zbl 0628.62011

[25] GURNEY, M. et JEWETT, R. S. (1975) Constructing orthogonal replications for standard errors. Journal of the American Statistical Association 70, 819-821. · Zbl 0322.62014

[26] HALL, P. G. (2003) A short prehistory of the bootstrap. Statistical Science 18, 158-167. · Zbl 1331.62018

[27] HAMPEL, F. R., RONCHETTI, E. M., ROUSSEEUW, P. J. et STAHEL, W. A. (1986) Robust Statistics : The Approach Based on Influence Functions. New York : Wiley. · Zbl 0593.62027

[28] HARTIGAN, J. A. (1969) Using subsample values as typical values. Journal of the American Statistical Association 64, 1303-1317.

[29] HUBER, P. J. (1981) Robust Statistics. New York : Wiley. · Zbl 0536.62025

[30] JUDKINS, D. R. (1990) Fay's method of variance estimation. Journal of Official Statistics 6, 223-239.

[31] KIM, J. K. et SITTER, R. R. (2003) Efficient replication variance estimation for two-phase sampling. Statistica Sinica 13, 641-653. · Zbl 1028.62007

[32] KOVAR, J. G., RAO, J. N. K. et WU, C. F. J. (1988) Bootstrap and other methods to measure errors in survey estimates. Canadian Journal of Statistics 16, 25-45. · Zbl 0663.62018

[33] KREWSKI, D. et RAO, J. N. K. (1981) Inference from stratified samples : Properties of the linearization, jackknife and balanced repeated replication methods. Annals of Statistics 9, 1010-1019. · Zbl 0474.62013

[34] LAHIRI, P. (2003) On the impact of bootstrap in survey sampling and small-area estimation. Statistical Science 18, 199-210. · Zbl 1331.62076

[35] MCCARTHY, P. J. (1969) Pseudo-replication : Half samples. Review of the Interational Statistics Institute 37, 239-264. · Zbl 0186.53001

[36] MCCARTHY, P. J. et SNOWDEN, C. B. (1985) The bootstrap and finite population sampling. Vital and Health Statistics 2, 2-95.

[37] MILLER, R. G. (1974) The jackknife - A review. Biometrika 61, 1-15. · Zbl 0275.62035

[38] NIELSEN, S. F. (2003) Proper and improper multiple imputation (with discussion). International Statistical Review 71 , 593-627. · Zbl 1114.62323

[39] POLITIS, D. N., ROMANO, J. P. et WOLF, M. (1999) Subsampling. New York : Springer-Verlag. · Zbl 0931.62035

[40] PRESNELL, B. et BOOTH, J. G. (1994) Resampling methods for sample surveys. Technical Report 470, Department of Statistics, University of Florida, Gainesville.

[41] QUENOUILLE, M. H. (1949a) Approximate tests of correlation in time-series. Journal of the Royal Statistical Society, Series B 11, 68-84. · Zbl 0035.09201

[42] QUENOUILLE, M. H. (1949b) Notes on bias in estimation. Biometrika 43, 353-360. · Zbl 0074.14003

[43] RAO, J. N. K. (1988) Variance estimation in sample surveys. In Handbook of Statistics, Volume 6, ed. P. R. K. amd C. R. Rao, pp. 427-447. Amsterdam : Elsevier Science.

[44] RAO, J. N. K. et SHAO, J. (1992) Jackknife variance estimation with survey data under hot deck imputation. Biometrika 79, 811-822. · Zbl 0764.62008

[45] RAO, J. N. K. et SHAO, J. (1996) On balanced half-sample variance estimation in stratified random sampling. Journal of the American Statistical Association 91 , 343-348. · Zbl 0871.62013

[46] RAO, J. N. K. et SHAO, J. (1999) Modified balanced repeated replication for complex survey data. Biometrika 86, 403-415. · Zbl 0933.62010

[47] RAO, J. N. K. et WU, C. F. J. (1985) Inference from stratified samples : Second-order analysis of three methods for nonlinear statistics. Journal of the American Statistical Association 80, 620-630. · Zbl 0603.62009

[48] RAO, J. N. K. et WU, C. F. J. (1988) Resampling inference with complex survey data. Journal of the American Statistical Association 83, 231-241. · Zbl 0654.62015

[49] RAO, J. N. K., WU, C. F. J. et YUE, K. (1992) Some recent work on resampling methods for complex surveys. Survey Methodology 18, 209-217.

[50] RENFER, J.-P. (2001) Description and process of the Household and Budget Survey of 1998 (HBS 1998). Swiss Federal Statistical Office. 1-19.

[51] RUBIN, D. B. (1987) Multiple Imputation for Nonresponse in Surveys. New York : Wiley. · Zbl 1070.62007

[52] RUBIN, D. B. (1996) Multiple imputation after 18+ years. Journal of the American Statistical Association 91, 473-489. with discussion, 507-515, and rejoinder, 515-517. · Zbl 0869.62014

[53] RUST, K. F. et RAO, J. N. K. (1996) Variance estimation for complex surveys using replication techniques. Statistical Methods in Medical Research 5, 283-310.

[54] SHAO, J. (2003) Impact of the bootstrap on sample surveys. Statistical Science 18, 191-198. · Zbl 1331.62078

[55] SHAO, J., CHEN, Y. et CHEN, Y. (1998) Balanced repeated replication for stratified multistage survey data under imputation. Journal of the American Statistical Association 93, 819-831. · Zbl 0947.62010

[56] SHAO, J. et RAO, J. N. K. (1994) Standard errors for low income proportions estimated from stratified multistage samples. Sankyā B 55, 393-414. · Zbl 0806.62100

[57] SHAO, J. et SITTER, R. R. (1996) Bootstrap for imputed survey data. Journal of the American Statistical Association 91, 1278-1288. · Zbl 0880.62011

[58] SHAO, J. et TU, D. (1995) The Jackknife and Bootstrap. New York : Springer-Verlag. · Zbl 0947.62501

[59] SHAO, J. et WU, C. F. J. (1989) A general theory for jackknife variance estimation. Annals of Statistics 17, 1176-1197. · Zbl 0684.62034

[60] SHAO, J. et WU, C. F. J. (1992) Asymptotic properties of the balanced repeated replication method for sample quantiles. Annals of Statistics 20, 1571-1593. · Zbl 0766.62005

[61] SITTER, R. R. (1992a) A resampling procedure for complex survey data. Journal of the American Statistical Association 87, 755-765. · Zbl 0760.62013

[62] SITTER, R. R. (1992b) Comparing three bootstrap methods for survey data. Canadian Journal of Statistics 20, 135-154. · Zbl 0753.62006

[63] SITTER, R. R. (1993) Balanced repeated replications based on orthogonal multi-arrays. Biometrika 80, 211-221. · Zbl 0769.62054

[64] TUKEY, J. W. (1958) Bias and confidence in not quite large samples (abstract). Annals of Mathematical Statistics 29, 614.

[65] WOLTER, K. M. (1985) Introduction to Variance Estimation. New York : Springer-Verlag.

[66] WU, C. F. J. (1991) Balanced repeated replications based on mixed orthogonal arrays. Biometrika 78, 181-188.

[67] ZHANG, P. (2003) Multiple imputation : Theory and method. International Statistical Review 71, 581-592. · Zbl 1114.65304