Complex Survey Data Analysis in Population-Based Epidemiology

Asst. Prof. Dr., Mehmet Emin ARAYICI

yaz
yayınları

Complex Survey Data Analysis in Population-Based Epidemiology

Asst. Prof. Dr., Mehmet Emin ARAYICI

yaz yayınları

2025

yaz
yayınları

Complex Survey Data Analysis in Population-Based Epidemiology

Yazar: Asst. Prof. Dr., Mehmet Emin ARAYICI

E_ISBN 978-625-5838-04-9

Ekim 2025 – Afyonkarahisar

Dizgi/Mizanpaj: YAZ Yayınları

Kapak Tasarım: YAZ Yayınları

YAZ Yayınları. Yayıncı Sertifika No: 73086

M.İhtisas OSB Mah. 4A Cad. No:3/3
İscehisar/AFYONKARAHİSAR

www.yazyayinlari.com

yazyayinlari@gmail.com

info@yazyayinlari.com

COMPLEX SURVEY DATA ANALYSIS IN POPULATION-BASED EPIDEMIOLOGY

Mehmet Emin ARAYICI[1]

1. INTRODUCTION

Population-based epidemiology relies on surveys that sample individuals from large populations to estimate health outcomes and risk factors. Unlike experiments, surveys do not manipulate exposure; they are observational, making representativeness and unbiased estimation essential. Many surveys employ complex sampling designs rather than simple random sampling. These designs typically involve stratifying the population into subgroups, sampling clusters of households or individuals, and applying sampling weights to adjust for unequal selection probabilities and non-response (Iparragirre, Barrio, Aramendi, & Arostegui, 2022; Parsons, Wei, & Parker, 2013). The aim is to reduce cost and increase precision while ensuring coverage of key subpopulations. Because units selected within the same cluster tend to be more similar, and because strata have different sampling fractions, the resulting data are correlated and heteroskedastic; ignoring the design elements leads to biased estimates and underestimated variances (Murillo Fort & Guillén Estany, 1989; Sturgis, 2004). Consequently, specialized methods are required for analyzing complex survey data in epidemiological research.

[1] Asst. Prof. Dr., Dokuz Eylül University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, mehmet.e.arayici@gmail.com, ORCID: 0000-0002-0492-5129.

This chapter provides a comprehensive overview of complex survey data analysis for population-based epidemiology. It begins by explaining sampling strategies such as stratification and clustering, then describes how sampling weights and post-stratification adjustments are constructed. Techniques for variance estimation—including Taylor series linearization and replication methods—are explored, followed by guidance on weighted regression modelling and strategies for dealing with missing data.

2. SAMPLING STRATEGIES: STRATIFICATION AND CLUSTERING

2.1. Stratified Sampling

Stratification divides the population into strata—homogeneous subgroups defined by characteristics such as region, urban/rural status, or demographic categories—and samples are drawn separately from each stratum. Stratification offers several advantages (Wu, Thompson, Wu, & Thompson, 2020)(Levin & Kanza, 2014):

Variance reduction: sampling homogeneous strata reduces variability within strata and leads to more precise estimates, as all strata are represented.

Ensuring representation of key subgroups: oversampling important but rare groups (e.g., racial/ethnic minorities) enhances power to analyse differences across subgroups.

Operational efficiency: fieldwork can be organized by region or cluster, reducing travel costs.

Two primary sampling strategies within stratification are proportional allocation and disproportional (or unequal) allocation. In proportional allocation, the sample fraction within

each stratum equals the population fraction, which minimizes overall variance when population variances are similar. In disproportional allocation, sample sizes are increased in smaller or higher-variability strata, often requiring weight adjustments to produce unbiased national estimates (Imrey, Sobel, & Francis, 1979; Yanagawa & Wakimoto, 1972).

Calculating stratum weights: For a stratified sample to produce national estimates, weights must reflect the population proportion of each stratum. The Micronutrient Survey Manual shows that stratum weights may be computed by dividing the population number by the sample number or by dividing the percentage of population by the percentage of sampled units; the selected method should be applied consistently across strata. Without weighting, estimates overrepresent heavily sampled strata and underrepresent lightly sampled strata, leading to biased national estimates (Kulas, Robinson, Smith, & Kellar, 2018; Schmidt et al., 2011).

2.2. Cluster Sampling

Cluster sampling selects clusters (primary sampling units or PSUs) such as villages, census blocks, or hospitals, then samples households or individuals within clusters. It is frequently combined with stratification to create a multi-stage design. Cluster sampling reduces fieldwork costs because interviewers travel to fewer locations, but it introduces intracluster correlation—people within the same cluster tend to be more similar than those across clusters—leading to increased sampling variance (Krenzke & Haung, 2014; Murphy & Chesnut, 2018). The design effect (DEFF) quantifies the inflation of variance due to clustering and weighting; it depends on the intracluster correlation coefficient (ICC) and average cluster size. Ignoring clustering yields standard errors that are too small and confidence

intervals that are too narrow, potentially leading to incorrect inferences (Bell et al., 2012).

Two-stage cluster sampling is common. In the Demographic and Health Surveys (DHS), a two-stage probability sample is drawn: first, PSUs stratified by region and urban/rural area are selected with probability proportional to size (PPS); second, households within PSUs are sampled systematically with equal probability (Aliaga & Ren, 2006; Kalton et al., 2021). The combination of stratification and clustering yields self-weighting samples within strata while achieving cost efficiency. Similar two-stage designs are used in other national surveys like the U.S. National Health and Nutrition Examination Survey (NHANES) ("National Health and Nutrition Examination Survey (NHANES) - Health, United States," 2024) and the Türkiye Health Surveys (Arayici & Kose, 2025).

2.3.Multi-stage and Multiphase Sampling

Large population surveys may involve more than two stages. For example, NHANES selects PSUs (counties), then segments (census blocks), then households, and finally individuals; there are often subsamples for laboratory measurements or special studies. Multiphase sampling may also be used, where certain measurements are obtained only for a subsample; weights for subsamples must be adjusted accordingly ("NHANES Tutorials - Weighting Module," n.d.-a).

3. SAMPLE WEIGHTING AND POST-STRATIFICATION

3.1. Purpose of Sampling Weights

Sampling weights are fundamental to complex survey analysis. They serve three purposes:

1. **Adjusting for unequal selection probabilities:**When the sampling design oversamples certain strata or clusters, the probability of selection varies. The base weight, defined as the inverse of the probability of selection, corrects this unequal probability (Bartosińska, 2012).
2. **Correcting for non-response:** Some sampled households or individuals do not participate. Non-response adjustments inflate weights of respondents within adjustment cells to compensate for missing units (Singh, Ganesh, & Lin, 2013).
3. **Post-stratification and calibration:** Final weights are calibrated to known population totals (e.g., by age, sex, or region) to correct residual bias and align sample distributions with census counts (Triveni, Danish, & Tawiah, 2024).

Ignoring weights can bias estimates because unweighted analyses treat each observation as equally representative. Inverse probability weighting reduces this bias but may increase variance if weights vary widely (Bell et al., 2012). The design effect accounts for this variance inflation.

3.2. Construction of Survey Weights

Survey weights typically evolve through a series of adjustments, which may vary by survey:

Base (design) weights: computed as the reciprocal of the selection probability. For example, NHANES (“NHANES Tutorials - Weighting Module,” n.d.-b) base weights combine the reciprocal of selecting the PSU, segment, household, and person.

Non-response adjustments: weights are multiplied by the inverse of the response rate within adjustment cells defined by demographics or design variables (Ezzati & Khare, 2002).

Post-stratification or raking: weights are calibrated to external population totals using demographic variables. NHANES calibrates its weights to estimates from the American Community Survey, replacing previous calibrations to the Current Population Survey to improve reliability for Asian populations. DHS computes separate household and individual weights (for women and men) by multiplying the inverse selection probability and the inverse of the household response rate; additional biomarker weights adjust for subsampling (Caughey et al., 2020).

3.3.Finite Population Correction and Replicate Weights

When a large fraction of the population is sampled (as may occur in small strata), a finite population correction (FPC) reduces variance estimates. FPC equals $(N-n)/(N-1)$where $N$is the population size and $n$the sample size; it reflects the reduction in variability because units cannot be resampled. Many software packages allow specifying FPC to improve accuracy ("Survey Data Analysis with R," n.d.). Replicate weights—sets of weights that replicate the sampling design across many replicates—provide an alternative to specifying strata and PSU variables. Balanced repeated replication (BRR) and jackknife replicate weights are common and are provided for surveys like the Current Population Survey. They replace PSU and strata information and allow straightforward estimation of variances

4. VARIANCE ESTIMATION TECHNIQUES

Accurate variance estimation is vital for valid inference in complex surveys. Standard errors computed under simple random sampling are biased when strata, clustering, and weights are ignored (Bell et al., 2012). Two major approaches are used: Taylor series linearization and replication methods.

4.1. Taylor Series Linearization

Taylor series linearization approximates the variance of a complex estimator by expanding it as a linear function around the population mean, then computing the variance of the linearized estimator. It is computationally efficient and generalizes well to regression coefficients and nonlinear statistics (Goodwin & Thompson, 2001; Hammer, Shin, & Porcellini, 2003). The NCES guide notes that Taylor linearization uses design variables (strata and PSU) and does not require replicate weights; it is available in software such as SAS PROC SURVEYREG ("SAS Help Center: Overview: SURVEYREG Procedure," n.d.), Stata svy commands ("How Do I Use the Stata Survey (Svy) Commands? | Stata FAQ," n.d.), SPSS Complex Samples ("Complex Samples - IBM SPSS Statistics," n.d.), and R's (Lumley, Gao, Schneider, & Lumley", 2025) survey package (So et al., 2020). For example, the variance of a mean $\hat{\mu} = \sum_i \quad w_i y_i / \sum_i \quad w_i$ can be linearized by approximating it with a ratio estimator and computing the variance of the numerator and denominator using design information. However, linearization requires deriving specific formulas for each estimator, which may be challenging for complex statistics like quantiles or poverty indices.

4.2. Replication Methods

Replication methods estimate variance by repeatedly re-computing the statistic on multiple replicate samples that mimic the complex design. The sampling variability across replicates approximates the variance of the full-sample estimator (Mukhopadhyay, An, Tobias, & Watts, 2008; Rust & Rao, 1996). Common methods include:

Balanced Repeated Replication (BRR): When the sample design contains strata with exactly two PSUs per stratum, BRR constructs replicates by retaining one PSU from each stratum according to a Hadamard matrix pattern and multiplying

its weight by two. The variance is calculated as the average squared deviation of replicate estimates from the full-sample estimate. Fay's BRR is a generalization that multiplies weights by factors ρ and $2-\rho$ rather than omitting PSUs, improving stability (Parsad & Gupta, 2007).

Jackknife Repeated Replication (JRR): The survey is divided into groups (often PSUs or clusters). Each replicate deletes one group and re-weights the remaining units. The variance is the scaled average of squared deviations of replicate estimates from the full estimate. JRR works with more than two PSUs per stratum and is widely used in surveys where replicate weights are provided (Frankel, 2014).

Bootstrap methods: Resampling of PSUs or clusters with replacement creates bootstrap replicates; weights are rescaled to preserve total population size. Bootstrapping is flexible and works for nonlinear statistics but may be computationally intensive (Chen & Sadler, 2010).

Replication methods have advantages when linearization formulas are difficult or when replicate weights are available from data producers. They also provide robust variance estimates for complex estimators. However, they require careful specification of replicate weights and may be sensitive to the number of replicates.

4.3. Design Effects and Intracluster Correlation

The design effect (DEFF) quantifies the inflation (or reduction) in variance due to the complex design relative to simple random sampling. It is defined as $\text{DEFF} = \text{var}_{\text{complex}} / \text{var}_{\text{srs}}$. Stratification typically reduces variance (DEFF<1) because homogeneous strata yield more precise estimates, whereas clustering increases variance (DEFF>1) due to correlated observations Survey weighting may reduce bias but can increase variance when weights have high variability. The intracluster

correlation coefficient (ICC) measures the similarity of units within clusters and drives the design effect; high ICC and large cluster size lead to large DEFF. Software packages often report the design effect for estimates, facilitating understanding of the efficiency of the design (Cohen, 2002; Kalton & Brick, 2005).

5. WEIGHTED REGRESSION MODELLING

Regression modelling in epidemiology often aims to relate exposures to health outcomes, adjusting for confounders. In complex surveys, regression coefficients must be estimated using design-appropriate methods to obtain unbiased estimates and correct standard errors (Daberkow, 1984).

In weighted least squares (WLS) regression, coefficients are estimated by minimizing the sum of squared residuals weighted by the sampling weights. The survey package in R implements WLS via svyglm(); Stata uses svy: regress; SPSS uses CSGLM (Complex Samples General Linear Models). Standard errors are estimated using Taylor linearization or replication. For logistic regression, the pseudo maximum likelihood estimator multiplies the log-likelihood by sampling weights, producing consistent parameter estimates; standard errors are obtained using the Huber–White sandwich estimator or design-based variance estimation. Proper specification of weights and design variables is critical; ignoring sampling weights can bias regression coefficients and understate standard errors (Bell et al., 2012; Blahut & Dayton, 2004).

An example is NHANES logistic regression for diabetes prevalence. Suppose we model diabetes status as a function of age, BMI, and sex. The design is specified by strata, PSU, and weights. In R: library(survey) design <- svydesign(ids=~psu, strata=~stratum, weights=~weight, data=nhanes) model <- svyglm(diabetes ~ age + bmi + sex, design=design,

family=quasibinomial()) The resulting coefficients estimate population log-odds ratios, and the svyglm function uses a sandwich variance estimator that accounts for clustering and stratification (Dey et al., 2025).

Outliers in survey data may correspond to influential clusters or misreported values. Weighted robust regression methods, such as M-estimators and GM-estimators, can reduce the influence of extreme observations. The robsurvey package extends the survey package by implementing robust estimators for complex survey data (Dehnel, 2016; Nugroho, Wardhani, Fernandes, & Solimun, 2020). It emphasises that different modes of inference—design-based, model-based and compound—reflect how much confidence is placed in model assumptions and that robust methods can mitigate sensitivity to outliers.

6. DEALING WITH MISSING DATA IN COMPLEX SURVEYS

Missing data are ubiquitous in epidemiological surveys because respondents may skip questions or units may drop out. Complex survey designs exacerbate the problem because missingness may be related to inclusion probabilities or design variables (Ghosh & Pahwa, 2008). Ignoring missingness can reduce the effective sample size and bias estimates.

Several strategies exist for handling missing data:

1. **Complete-case (listwise) deletion:** excludes cases with missing values. Although easy to implement, this strategy can bias estimates if the missingness mechanism is not completely at random; it also reduces sample size (Kellermann, Trevathan, & Kromrey, 2016).
2. **Inverse Probability Weighting (IPW):** uses weights proportional to the inverse probability of responding. A

logistic regression model predicts response probability using available variables (including design variables), and the inverse predicted probabilities serve as weights (Dey et al., 2025). IPW can correct for nonresponse bias when the response model is correctly specified, but extreme weights can inflate variance.

3. **Single Imputation (SI):** replaces missing values with predicted values or mean values; this tends to underestimate variance because it ignores imputation uncertainty (McMillan, 2013).
4. **Multiple Imputation (MI):** creates multiple complete datasets by replacing missing values with draws from an imputation model that includes predictors of missingness and the survey design variables. Estimates are computed on each imputed dataset, and results are combined using Rubin's rules. The logistic regression review notes that MI is popular for complex surveys; the pooled estimate averages across imputations, and the variance combines within- and between-imputation variability. MI models should include sampling weights, strata, and clusters as predictors or via dedicated survey imputation procedures (Zhang et al., 2023). Sequential regression multiple imputation (SRMI), also called fully conditional specification, is particularly suited for large surveys; each variable with missingness is imputed sequentially using regression models and includes design variables (Zhang et al., 2023).

Including design information in the imputation model is crucial. The skip-pattern imputation study emphasises that sampling weights, strata, and clustering should be included as predictors in the imputation model to preserve design-based inference; failing to include them can distort variance estimation

(Zhang et al., 2023). After imputation, the final analysis uses the survey design specification and combines estimates across imputations. A practical guide to multiple imputation in large surveys found that combining IPW and MI can improve accuracy and reduce bias compared with complete-case analysis; however, the nonresponse model may have limited predictive power and the selection of predictors is important (He, Zaslavsky, Harrington, Catalano, & Landrum, 2010).

7. SOFTWARE TOOLS FOR SURVEY ANALYSIS

R survey and Related Packages: The R language provides comprehensive tools for complex survey analysis. The survey package, developed by Thomas Lumley, supports specifying survey designs via svydesign(), which takes arguments for PSU (ids), strata, weights, and FPC (Lumley et al., 2025). It implements descriptive statistics, ratio estimation, quantiles, and regression models using linearization or replication. Weighted regression is implemented in svyglm() and svycoxph() for survival analysis; variance estimation options include linearization and replicate weights. The package also supports calibration and raking (via svycalibrate) and provides functions to compute design effects and intracluster correlation. The robsurvey package adds robust regression estimators for complex surveys (*Outline*, n.d.). For replication methods, the srvyr and tidyverse interfaces provide tidy syntax and functions for BRR and JRR; replicate weights can be created with as.svrepdesign().

Stata: svy Commands: Stata contains a suite of survey commands (svy) that incorporate sampling weights, stratification, and clustering **(Kolenikov, 2008; Winter, 2002)**. Users specify the design using svyset, providing the weight variable (pweight), stratification (strata), PSU (psu), FPC, and replicate weights if available. Stata implements descriptive statistics (svy: mean, svy:

proportion), linear regression (svy: regress), generalized linear models (svy: glm), logistic regression (svy: logistic), and survival analysis. The UCLA guide explains that probability weights in Stata correspond to final survey weights that include all adjustments; replicate weights can be used when PSU or strata variables are unavailable. Stata computes standard errors using Taylor linearization by default and provides options for Jackknife, BRR, and bootstrap variance estimation.

SPSS Complex Samples: IBM SPSS Statistics offers the Complex Samples modüle **("Complex Samples - IBM SPSS Statistics," n.d.)**, which incorporates sample design information into analysis. The module provides procedures for:

Complex Samples Descriptives (CSDESCRIPTIVES): computes means, sums, ratios and design effects, with variance estimation based on the sample design (equal probability or PPS) and sampling with or without replacement. Complex Samples Tabulate (CSTABULATE): generates frequency tables, cross-tabulations, and design effects. Complex Samples General Linear Models (CSGLM): performs linear regression, ANOVA, and ANCOVA using sampling weights and design variables; outputs parameter estimates, standard errors, and tests of hypotheses. Complex Samples Logistic Regression (CSLOGISTIC): fits logistic regression models with sampling weights, providing coefficient estimates, standard errors, odds ratios, design effects, and Wald tests.

The module allows specifying sampling plans, including equal probability designs, PPS sampling, and designs with or without replacement. Users can handle missing data via listwise deletion or imputation within the module. Other software includes SAS **("SAS Help Center: Overview: SURVEYREG Procedure," n.d.)** (PROC SURVEYMEANS, PROC SURVEYREG, PROC SURVEYLOGISTIC), SUDAAN, and

specialized packages like convey for poverty measures in R. The guidelines provided by the NCES emphasise using software that supports design-based variance estimation to avoid underestimating variances.

8. LIMITATIONS AND CHALLENGES

Despite the availability of sophisticated methods and software, complex survey analysis faces several limitations.

1. **Design mis-specification:** Analysts must correctly specify strata, PSU, weights, and FPC. Mis-specification leads to biased variance estimation. In secondary data analyses, documentation may be incomplete, and replicate weights may be unavailable.
2. **Extreme or highly variable weights:** Large variation in weights can inflate variance and produce unstable estimates. Weight trimming or smoothing may reduce variance at the cost of introducing bias.
3. **Non-response and missing data:** Response rates may be low, and missing data may correlate with design variables. IPW and MI rely on correct models; misspecification can bias results. Complex skip patterns and skip-condition variables further complicate imputation.
4. **Measurement error and recall bias:** Self-reported outcomes (e.g., dietary intake, physical activity) may be measured with error. Complex survey analysis does not inherently correct measurement error; additional methods such as validation substudies are needed.
5. **Computational resources:** Large surveys with replicate weights require substantial memory and computing time, especially when using bootstrap methods.

6. **Causal inference:** Complex survey data are observational; causal interpretations require careful consideration of confounding and selection bias. Incorporating design variables into causal models remains an area of methodological research.

Advances in survey methodology aim to address these challenges and enhance the utility of complex surveys for epidemiology. Linking survey data with administrative records, electronic health records or wearable devices can enrich analyses and reduce respondent burden. However, linkage introduces privacy concerns and requires harmonizing sampling weights. Surveys are adopting adaptive designs that modify sampling probabilities in real time based on paradata (e.g., response propensities), improving efficiency and reducing nonresponse bias. Responsive design may allocate more resources to strata with lower response rates. Combining design-based weights with flexible modelling (e.g., hierarchical models, Bayesian inference) can improve estimates for small domains and incorporate prior information. The tension between design-based and model-based approaches is being addressed through compound inference methods that balance the strengths of both. Machine learning algorithms can model complex relationships among variables to improve nonresponse adjustment, calibration, and imputation. Ensuring interpretability and maintaining design-based properties remain active research areas. New variance estimation techniques aim to handle high-dimensional data, complex estimators, and robust measures of association. For example, robust regression and influence function approaches provide resilience against outliers and model misspecification. To facilitate secondary analysis, survey organizations are increasingly providing detailed metadata, design variables, replicate weights, and user guides (e.g., DHS, NHANES).

Continued improvements in accessibility will reduce errors and broaden the use of complex survey data.

9. CONCLUSIONS

Complex survey designs are indispensable for collecting representative epidemiological data at national and subnational levels. By combining stratification, multi-stage cluster sampling, and weighting, surveys like NHANES, DHS, and THS efficiently produce data to monitor population health. Analysis of such data requires explicit recognition of the sampling design: weights adjust for unequal selection and nonresponse, and variance estimation must account for stratification and clustering. Taylor series linearization and replication methods provide reliable standard errors; design effects inform sample efficiency; and weighted regression models yield valid associations. Handling missing data through IPW and multiple imputation and using specialized software tools are essential for robust analysis. While challenges remain—such as design mis-specification, extreme weights, and model misspecification—ongoing methodological advances promise to strengthen the integration of complex survey data into population-based epidemiology. Careful application of design-based principles and awareness of survey documentation will ensure that epidemiological inferences reflect the populations the surveys are intended to represent.

REFERENCES

Aliaga, A., & Ren, R. (2006, July 1). *Optimal sample sizes for two-stage cluster sampling in demographic and health surveys*. Retrieved from https://www.semanticscholar.org/paper/Optimal-sample-sizes-for-two-stage-cluster-sampling-Aliaga-Ren/99d744d0877d680d01de922398cc23bf22afd898?utm_source=consensus

Arayici, M. E., & Kose, A. (2025). Prevalence of Alzheimer's Disease and Cardiometabolic Multimorbidity in Older Adults Aged 60 and above in a Large-Scale Representative Sample in Türkiye: A Nationwide Population-Based Cross-Sectional Study. *Journal of Epidemiology and Global Health*, *15*(1), 86. https://doi.org/10.1007/s44197-025-00435-5

Bartosińska, D. (2012). *Statistical Inference from Complex Sample with SAS on the Example of Household Budget Surveys*. Retrieved from https://www.semanticscholar.org/paper/Statistical-Inference-from-Complex-Sample-with-SAS-Bartosi%C5%84ska/c971345e4b162ef7a4def8d63da9e7045402fe79?utm_source=consensus

Bell, B. A., Onwuegbuzie, A. J., Ferron, J. M., Jiao, Q. G., Hibbard, S. T., & Kromrey, J. D. (2012). Use of Design Effects and Sample Weights in Complex Health Survey Data: A Review of Published Articles Using Data From 3 Commonly Used Adolescent Health Surveys. *American Journal of Public Health*, *102*(7), 1399–1405. https://doi.org/10.2105/AJPH.2011.300398

Blahut, S., & Dayton, C. (2004). *Latent class logistic regression with complex sample survey data*. Retrieved from https://www.semanticscholar.org/paper/Latent-class-

logistic-regression-with-complex-data-Blahut-Dayton/26b6724a2a94568161b5588878f2e365d7255e73?utm_source=consensus

Caughey, D., Berinsky, A. J., Chatfield, S., Hartman, E., Schickler, E., & Sekhon, J. S. (2020, October 22). *Target Estimation and Adjustment Weighting for Survey Nonresponse and Sampling Bias*. Cambridge University press. https://doi.org/10.1017/9781108879217

Chen, H., & Sadler, A. (2010). *Comparison of Variance Estimation Methods Using PPI Data October 2010*. Retrieved from https://www.semanticscholar.org/paper/Comparison-of-Variance-Estimation-Methods-Using-PPI-Chen-Sadler/83ac79a55b74d027d44ff06d151d042f70aac2b3?utm_source=consensus

Cohen, S. (2002). *Design Effect Variation In The National Medical Care Expenditure Survey*. Retrieved from https://www.semanticscholar.org/paper/DESIGN-EFFECT-VARIATION-IN-THE-NATIONAL-MEDICAL-Cohen/becefc9769b08f37697b29b19b96c6592aef3b13?utm_source=consensus

Complex Samples—IBM SPSS Statistics. (n.d.). Retrieved October 9, 2025, from https://www.ibm.com/products/spss-statistics/complex-samples

Daberkow, S. (1984). *Regression Analysis With Complex Survey Data: A Comparison Of Estimation Techniques*. Retrieved from https://www.semanticscholar.org/paper/Regression-Analysis-With-Complex-Survey-Data%3A-A-Of-

Daberkow/0aff247815b979e3bfc9fb2ecfc0a8f647abfd42?utm_source=consensus

Dehnel, G. (2016). M-Estimators in Business Statistics[1]. *Statistics in Transition New Series*, *17*(4), 749–762. https://doi.org/10.21307/stattrans-2016-050

Dey, D., Haque, Md. S., Islam, Md. M., Aishi, U. I., Shammy, S. S., Mayen, Md. S. A., … Uddin, Md. J. (2025). The proper application of logistic regression model in complex survey data: A systematic review. *BMC Medical Research Methodology*, *25*, 15. https://doi.org/10.1186/s12874-024-02454-5

Ezzati, T., & Khare, M. (2002). *Nonresponse Aimustmen'i~ In A National Health Survey*. Retrieved from https://www.semanticscholar.org/paper/NONRESPONSE-AIMUSTMEN'I~-IN-A-NATIONAL-HEALTH-Ezzati-Khare/2c576bf8eec8749e35e87a6fb8345baa04e0c770?utm_source=consensus

Frankel, M. R. (2014, September 29). *Resampling Procedures for Sample Surveys* (R. S. Kenett, N. T. Longford, W. W. Piegorsch, & F. Ruggeri, Eds.). Wiley. https://doi.org/10.1002/9781118445112.stat05043

Ghosh, S., & Pahwa, P. (2008). *Assessing Bias Associated With Missing Data from Joint Canada/U.S. Survey of Health: An Application*. Retrieved from https://www.semanticscholar.org/paper/Assessing-Bias-Associated-With-Missing-Data-from-of-Ghosh-Pahwa/c59a5e2e7e5f52baa100b7e26c41c61c0a6920b3?utm_source=consensus

Goodwin, R., & Thompson, K. J. (2001). *Using Taylors Linearization Technique in StEPS to Estimate Variances*

for Non-Linear Survey Estimators. Retrieved from https://www.semanticscholar.org/paper/Using-Taylors-Linearization-Technique-in-StEPS-to-Goodwin-Thompson/620f28e99ef18a32ca084254a2db326362ccf0e0?utm_source=consensus

Hammer, H., Shin, H.-C., & Porcellini, L. (2003). *A Comparison of Taylor Series and JK1 Resampling Methods for Variance Estimation*. Retrieved from https://www.semanticscholar.org/paper/A-Comparison-of-Taylor-Series-and-JK1-Resampling-Hammer-Shin/8de15f80a7c0b12df58f6cc684701faf634a8b4f?utm_source=consensus

He, Y., Zaslavsky, A. M., Harrington, D. P., Catalano, P., & Landrum, M. B. (2010). Multiple Imputation in a Large-Scale Complex Survey: A Practical Guide. *Statistical Methods in Medical Research, 19*(6), 653–670. https://doi.org/10.1177/0962280208101273

How do I use the Stata survey (svy) commands? | Stata FAQ. (n.d.). Retrieved October 9, 2025, from https://stats.oarc.ucla.edu/stata/faq/how-do-i-use-the-stata-survey-svy-commands/

Imrey, P. B., Sobel, E., & Francis, M. E. (1979). Analysis of categorical data obtained by stratified random sampling. *Communications in Statistics - Theory and Methods*, *8*(7), 653–670. https://doi.org/10.1080/03610927908827790

Iparragirre, A., Barrio, I., Aramendi, J., & Arostegui, I. (2022). Estimation of cut-off points under complex-sampling design data. *SORT. Statistics and Operations Research Transactions,* (46), 137–158. https://doi.org/10.2436/20.8080.02.121

Kalton, G., & Brick, J. (2005). *Estimating components of design effects for use in sample design*. Retrieved from https://www.semanticscholar.org/paper/Estimating-components-of-design-effects-for-use-in-Kalton-Brick/6541311231bcc7b257c75163ffb55cb8be39e81d?utm_source=consensus

Kalton, G., Flores Cervantes, I., Arieira, C., Kwanisai, M., Radin, E., Saito, S., … Stupp, P. (2021). Dealing With Inaccurate Measures Of Size In Two-Stage Probability Proportional To Size Sample Designs: Applications In African Household Surveys. *Journal of Survey Statistics and Methodology*, *9*(5), 1035–1049. https://doi.org/10.1093/jssam/smaa020

Kellermann, A. P., Trevathan, D., & Kromrey, J. (2016). *Missing Data and Complex Sample Surveys Using SAS ®: The Impact of Listwise Deletion vs . Multiple Imputation on Point and Interval Estimates when Data are MCAR and MAR*. Retrieved from https://consensus.app/papers/missing-data-and-complex-sample-surveys-using-sas-%C2%AE-the-kromrey-trevathan/a732020a80fc592a8283140c12e75551/

Kolenikov, S. (2008). *Survey bootstrap and bootstrap weights*. Retrieved from https://consensus.app/papers/survey-bootstrap-and-bootstrap-weights-kolenikov/c9137b231e35564babbe38a28de046bb/

Krenzke, T., & Haung, W.-C. (2014). *Revisiting Nested Stratification of Primary Sampling Units*. Retrieved from https://www.semanticscholar.org/paper/Revisiting-Nested-Stratification-of-Primary-Units-Krenzke-Haung/7e6164c72d4754f3d880ed4f5d322b50babe7526?utm_source=consensus

Kulas, J. T., Robinson, D. H., Smith, J. A., & Kellar, D. Z. (2018). Post-Stratification Weighting in Organizational Surveys: A Cross-Disciplinary Tutorial. *Human Resource Management*, *57*(2), 419–436. https://doi.org/10.1002/hrm.21796

Levin, R., & Kanza, Y. (2014). Stratified-sampling over social networks using mapreduce. *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*, 863–874. https://doi.org/10.1145/2588555.2588577

Lumley, T., Gao, P., Schneider, B., & Lumley", "Thomas. (2025). *survey: Analysis of Complex Survey Samples*. Retrieved from https://cran.r-project.org/web/packages/survey/index.html

McMillan, S. T. (2013). *Comparison of Imputation Methods in the Survey of Income and Program Participation*. Retrieved from https://www.semanticscholar.org/paper/Comparison-of-Imputation-Methods-in-the-Survey-of-McMillan/79bffcfec4f76c55fe9ce0b24c528c9f00f63e52?utm_source=consensus

Mukhopadhyay, P., An, A., Tobias, R., & Watts, D. (2008). *Try, Try Again: Replication-Based Variance Estimation Methods for Survey Data Analysis in SAS ® 9.2*. Retrieved from https://www.semanticscholar.org/paper/Try%2C-Try-Again%3A-Replication-Based-Variance-Methods-Mukhopadhyay-An/4e9ff130e5f6ad58366b6513915dda98a79d6eba?utm_source=consensus

Murillo Fort, C., & Guillén Estany, M. (1989). [Estimation of the variances of the variables in the Barcelona health survey].

Gaceta Sanitaria, *3*(12), 409–420. https://doi.org/10.1016/s0213-9111(89)70962-6

Murphy, P., & Chesnut, J. (2018). *A Comparison of Clustering Algorithms used for Multivariate Stratification of Primary Sampling Units ⁎.* Retrieved from https://www.semanticscholar.org/paper/A-Comparison-of-Clustering-Algorithms-used-for-of-%E2%88%97-Murphy-Chesnut/b176c7fd7a66e9809bd42e937a2e580762c39370?utm_source=consensus

National Health and Nutrition Examination Survey (NHANES)—Health, United States. (2024, July 29). Retrieved October 6, 2025, from https://www.cdc.gov/nchs/hus/sources-definitions/nhanes.htm

NHANES Tutorials—Weighting Module. (n.d.-a). Retrieved October 6, 2025, from https://wwwn.cdc.gov/nchs/nhanes/tutorials/weighting.aspx

NHANES Tutorials—Weighting Module. (n.d.-b). Retrieved October 6, 2025, from https://wwwn.cdc.gov/nchs/nhanes/tutorials/weighting.aspx

Nugroho, W. H., Wardhani, N. W. S., Fernandes, A. A. R., & Solimun, S. (2020). Robust Regression Analysis Study for Data with Outliers at Some Significance Levels. *Mathematics and Statistics*, *8*(4), 373–381. https://doi.org/10.13189/ms.2020.080401

Outline. (n.d.). Retrieved from https://cran.r-project.org/web/packages/robsurvey/vignettes/regression.html

Parsad, R., & Gupta, V. K. (2007). *Variance Estimation From Complex Surveys Using Balanced Repeated Replication*. Retrieved from https://www.semanticscholar.org/paper/VARIANCE-ESTIMATION-FROM-COMPLEX-SURVEYS-USING-Parsad-Gupta/4e2ec8458c5a4408837731bf9bc41aeee8412762?utm_source=consensus

Parsons, V., Wei, R., & Parker, J. (2013). *Evaluation of Model-based Methods in Analyzing Complex Survey Data: A Simulation Study using Multistage Complex Sampling on a Finite Population 1*. Retrieved from https://www.semanticscholar.org/paper/Evaluation-of-Model-based-Methods-in-Analyzing-A-on-Parsons-Wei/1d107647287365eefdabb3f7520b5403988e6c69?utm_source=consensus

Rust, K. F., & Rao, J. N. (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, *5*(3), 283–310. https://doi.org/10.1177/096228029600500305

SAS Help Center: Overview: Surveyreg Procedure. (n.d.). Retrieved October 9, 2025, from https://documentation.sas.com/doc/en/pgmsascdc/v_067/statug/statug_surveyreg_overview.htm?utm_source=chatgpt.com

Schmidt, C. O., Alte, D., Völzke, H., Sauer, S., Friedrich, N., & Valliant, R. (2011). Partial misspecification of survey design features sufficed to severely bias estimates of health-related outcomes. *Journal of Clinical Epidemiology*, *64*(4), 416–423. https://doi.org/10.1016/j.jclinepi.2010.04.019

Singh, A., Ganesh, N., & Lin, Y. (2013). *Improved Sampling Weight Calibration by Generalized Raking with Optimal Unbiased Modification*. Retrieved from https://www.semanticscholar.org/paper/Improved-Sampling-Weight-Calibration-by-Generalized-Singh-Ganesh/53fe6a1565f3284e6f5373018b01809cbfab755f?utm_source=consensus

So, H. Y., Oz, U. E., Griffith, L., Kirkland, S., Ma, J., Raina, P., … Wu, C. (2020, October 24). *Modelling Complex Survey Data Using R, SAS, SPSS and Stata: A Comparison Using CLSA Datasets*. arXiv. https://doi.org/10.48550/arXiv.2010.09879

Sturgis, P. (2004, October 1). *Analysing Complex Survey Data: Clustering, Stratification and Weights*. Retrieved from https://www.semanticscholar.org/paper/Analysing-Complex-Survey-Data%3A-Clustering%2C-and-Sturgis/b875ef9471165bd981d6507445c42e79c208a2d7?utm_source=consensus

Survey Data Analysis with R. (n.d.). Retrieved October 6, 2025, from https://stats.oarc.ucla.edu/r/seminars/survey-data-analysis-with-r/

Triveni, G. R. V., Danish, F., & Tawiah, K. (2024). Evolving techniques for enhanced estimation: A comprehensive survey of stratified sampling and post-stratification methods. *AIP Advances*, *14*(10), 100703. https://doi.org/10.1063/5.0193961

Winter, N. (2002). SVR: Stata module to compute estimates with survey replication (SVR) based standard errors. *Statistical Software Components*. Retrieved from https://consensus.app/papers/svr-stata-module-to-compute-estimates-with-survey-winter/98273ed0b804570ba1358585f1d818e8/

Wu, C., Thompson, M. E., Wu, C., & Thompson, M. E. (2020). *Stratified Sampling and Cluster Sampling*. 33–56. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-44246-0_3

Yanagawa, T., & Wakimoto, K. (1972). Estimation of some functional of the population distribution based on a stratified random sample. *Annals of the Institute of Statistical Mathematics*, *24*, 137–151. https://doi.org/10.1007/BF02479745

Zhang, G., He, Y., Cai, B., Moriarity, C., Shin, H.-C., Parsons, V., & Irimata, K. E. (2023). Multiple imputation of missing data with skip-pattern covariates: A comparison of alternative strategies. *Journal of Statistical Computation and Simulation*, *94*(7), 1543–1570. https://doi.org/10.1080/00949655.2023.2293124

Complex Survey Data Analysis in Population-Based Epidemiology

yaz
yayınları

YAZ Yayınları
M.İhtisas OSB Mah. 4A Cad. No:3/3
İscehisar / AFYONKARAHİSAR
Tel : (0 531) 880 92 99
yazyayinlari@gmail.com • www.yazyayinlari.com