Effects of calculating standard errors using estimation versus variance weights for 2004 National Nursing Home Survey

Rong Cai[1], Sarah Gousen[2], Iris Shimizu[2], Robin Remsburg[2]
[1]National Center for Health Statistics, CDC
[2]NCHS, CDC

Abstract

The design for the National Nursing Home Survey (NNHS in-house file) specifies two weights: an estimation weight to calculate estimates and a variance weight to determine sampling errors of the estimates. Estimation weights are the product of ratio adjustment factors and variance weights. Software restrictions in programs such as SUDAAN require two separate computations: one for the estimate and one for the standard error, a requirement which increases analytic complexity and can be a source of error. This paper discusses an investigation into the differences between sampling errors based on the two sets of weights. The effects of using estimation weights instead of variance weights to compute the sampling errors in analysis are evaluated; and instructions are offered to minimize erroneous conclusions when working with publicly released data, which includes only the estimation weight.

Keywords: Standard errors, Variance weights, Ratio adjustment

1. Introduction

Standard error calculation is a very important part of data analysis for health surveys like the 2004 National Nursing Home Survey. Standard errors are important for testing hypotheses and the formation of confidence intervals. For a fixed significance level, underestimated standard errors increase the likelihood of rejecting the null hypothesis, which is a Type I error if the hypothesis is true. The probability of committing a Type I error is denoted as significance level α. [1] Similarly, overstated standard errors increase the likelihood of failing to reject a null hypothesis when it actually is false. This is a Type II error. To be conservative, NCHS places a priority on minimizing Type I errors. On the other hand, standard errors that are excessively overstated (say by 10 percent) are not useful. Avoiding the increased complexity of separate computations for the estimate and for the standard error was one goal, while the priority to minimize Type I errors was a second. The conflict between these two goals motivated this investigation into the effect that a ratio adjusted estimation weight had on standard errors when compared to the unadjusted variance weight.

The NNHS survey collects data on nursing homes in the United States, on individuals residing in them, and on nursing assistants employed in them. The 2004 NNHS used a stratified two-stage probability design. The first-stage consists of a probability sample of nursing facilities. The sampling frame for the 2004 NNHS consisted of nursing homes (and their bed sizes) that were certified for Medicare and/or Medicaid; and non-certified nursing homes that were included in a computerized file of nursing homes obtained from a private vendor. The second-stage consists of a systematic sample of residents from each selected nursing facility; or of nursing assistants (for the resident and nursing assistant data, respectively). See NCHS [2] for more details of the 2004 NNHS design. Here, we focus our research on the first-stage facility level.

Two types of facility level statistics are produced in the 2004 NNHS; one is for variables that are not correlated with bed size (for example, the number of facilities certified by Medicare) and the other is for variables that are correlated with bed size (for example, the number of registered nurses employed in nursing homes). Two sets of weights are necessary to permit realistic estimates about facility attributes from a sample that was selected with probability proportional to facility bed counts. Each estimation weight is the product of three factors:

- the inverse of the facility selection probability
- an adjustment for facility non-response, and
- a ratio adjustment corresponding to each region r (r = 1, 2, 3, 4).

These weights are referred to as the estimation weights, in contrast to other weights, in the discussion below.

The differences between the weights for estimates correlated with bed size and those which are not lie in different computations of the non-response and ratio adjustments. Facility counts were used in the non-response and the ratio adjustments for the estimates which are not correlated with bed sizes while bed counts were used in those adjustments for the estimates which are correlated with bed size. The ratio adjustments (Y_r/Y_r') are designed to reduce the variances of the NNHS estimates by calibrating the estimates to account for the total population of nursing homes (or their beds) which are listed in the NNHS frame for 2004, the year in which the survey data were collected. The numerator Y_r of a ratio adjustment factor for each region is the total number of facilities (or their beds) listed for that region in the NNHS sampling frame while the denominator Y_r' is an estimate of the numerator based on sampling frame information for the NHHS sampled facilities.

Because of the ratio adjustment, the NNHS estimates for individual regions are actually ratio estimates of the form $Y_r\left(X_r'/Y_r'\right)$ where X_r' and Y_r' are each random variables. The X_r' is a non-ratio-adjusted estimate; that is, the weights for the numerator estimates X_r' are the products of only two factors: the inverse of the facility selection probabilitiy and an adjustment for non-response. These weights for X_r' are labeled as "variance weights" in the following. The Y_r' for the denominator was discussed above and its weights are the inverses of the facility selection probabilities. The estimates for other facility groups are sums of the regional ratio estimates.

Because NNHS uses a complex sample, variance calculations need to account for the sample design. Software used for analysis of data from complex samples includes SUDAAN [3], STATA, SPSS, and SAS survey procedures. Among them, SUDAAN's linearized Taylor series software is widely used within NCHS and is used in the analysis discussed in this paper. SUDAAN does not accept use of two weights in one WEIGHT statement. Because ratio adjustments are designed to reduce variances, the variances of estimates that are not ratio adjusted are most likely equal or larger than those of the corresponding ratio adjusted estimates. Hence, because it is deemed better to overstate than to understate variances in analysis, the variances of the NNHS estimates are approximated by calculating the variances for the corresponding non-ratio-adjusted estimates. That is, two SUDAAN runs are used, one using estimation weights to produce the ratio adjusted estimates and the other using variance weights to approximate the variances of those estimates. This requirement for two software runs may confuse analysts and contribute to errors. The use of only the estimation weights in the public use file to produce both estimates and their variances was viewed as a way to reduce confusion and minimize errors. (The variance weights are available upon request.) However, use of the estimation weights may over or underestimate the standard errors adopted as "official" and, thus, affect the accuracy of statistical results. This paper discusses an investigation into the differences between variances estimated by using estimation and variance weights, and recommends procedures for adjusting the results of statistical tests to account for under-statement of variances.

2. Notation and formulas:

Let n_r be the number of respondent sample facilities in region r (r = 1, 2, 3, 4). For the x-variable of interest and facility i (I = 1, 2, …, n_r) in region r, let

x_{ri} be the observed value of X,

w_{vri} = the variance weight, and

w_{Xeri}= R_{Xr} w_{vri} = the estimation weight for the X-variable where

R_{Xr} denotes the ratio adjustment factor for the X-variable in region r.

Then, the estimators of X (sum of x-values) for facilities in region r and the nation may be formulated as

$$\hat{X}_r = \sum_i w_{Xeri} x_{ri} \quad \text{and} \tag{1}$$

$$\hat{X} = \sum_r \hat{X}_r = \sum_r \sum_i R_{Xr} w_{vri} x_{ri} \ , \tag{2}$$

respectively. Let S_{Xv} and S_{Xe} denote the estimators for the standard errors of $\hat{X}$ which are calculated by using the variance and estimation weights, respectively. Also, let $C_X = S_{Xe} / S_{Xv}$, the ratio of the standard errors for the X variable.

Because the calibration ratio adjustments are defined by region, the R_{xr} is treated as a constant when calculating variance approximations for regional estimate $\hat{X}_r$ in (2) with estimation weights. That is, $S_{Xer} = R_{Xr} S_{Xvr}$ (r = 1, 2, 3, 4) so that $C_{Xr} = R_{Xr}$. For NNHS 2004, the four region adjustment factor values for X variables which are not related to bed size are R_1 = 1.025, R_2 = 0.977, R_3 = 0.993, and R_4 = 1.011. For X variables which are related to bed size are R_1 = 1.014, R_2 = 0983, R_3 = 0.999, and R_4 = 1.050. Because the 2004 nursing home sample was not stratified by region, there is no simple expression for the relationship between the calibration ratios and the S_{xe} and S_{Xv} for estimates about other facility groups.

3. Analytic methods

The effect of using estimation weights instead of variances weights to calculate variance estimates was conducted by comparing the 95% confidence intervals estimated with S_{Xe} and S_{Xv} , respectively. For the purposes of these evaluations, the S_{Xv} estimates were assumed to be the gold standards and, hence, the interval

$$\left(\hat{X} - 1.96S_{Xv},\ \hat{X} + 1.96S_{Xv}\right). \quad (3)$$

estimated with S_{Xv} was assumed to be the true 95% confidence interval for X. If the interval in (3) is contained entirely within the 95% confidence interval estimated with S_{Xe}, then the intervals estimated with S_{Xe} have more than a 95 % chance of including X and, hence, use of the S_{Xe} would give more conservative statistical test results than would use of S_{Xv}.

Let $S_{Xe} = C_X\ S_{Xv}$. The 95% confidence interval estimated by using S_{Xe} is

$$\left(\hat{X} - 1.96S_{Xe}, \hat{X} + 1.96S_{Xe}\right) \quad (4)$$

Rewriting this interval in terms of S_{Xv} instead of S_{Xe}, the confidence interval is actually

$$\left(\hat{X} - 1.96C_X S_{Xv},\ \hat{X} + 1.96C_X S_{Xv}\right). \quad (5)$$

Case 1: If $C_X > 1$, then interval (5) entirely contains the targeted interval $\left(\hat{X} - 1.96S_{Xv}, \hat{X} + 1.96S_{Xv}\right)$. In this case, the interval defined in (4) has more than a 95% chance of including X and the use of S_{Xe} will give more conservative statistical test results than use of S_{Xv}.

Case 2: If $C_X < 1$, then the actual interval (5) is a subset of the targeted interval ($\hat{X} - 1.96S_{Xv},\ \hat{X} + 1.96S_{Xv}$) in (3) so that it covers X less than 95% of the time. In order to increase the confidence level to 95%, one can enlarge the interval range by dividing the coefficients of S_{Xv} in the actual interval (5) by the minimum value of the C_X over all study variables as follows:

$$\left(\hat{X} - \frac{1.96C_X S_{Xv}}{\left[\min_Y\left(C_Y\right)\right]},\ \hat{X} + \frac{1.96C_X S_{Xv}}{\left[\min_Y\left(C_Y\right)\right]}\right), \quad (6)$$

where the study Xs are relabeled as Ys for the minimum. Rewriting interval (6) in terms of S_{Xe} instead of S_{Xv}, the interval becomes

$$\left(\hat{X} - \frac{1.96}{\left[\min_Y\left(C_Y\right)\right]}S_{Xe}, \hat{X} + \frac{1.96}{\left[\min_Y\left(C_Y\right)\right]}S_{Xe}\right). \quad (7)$$

This means, that to assure a minimum of 95% in confidence intervals or a minimum of 5% significance levels in hypothesis tests when using S_{Xe} in place of S_{Xv}, one should lower the α value from 5% to one (from a table for the standard normal distribution) which corresponds with coefficients $\pm\ 1.96\big/\left[\min_Y\left(C_Y\right)\right]$ of S_{Xe} in interval (7). See the end of Section 4 for an example of adjustments to the α value.

4. Results from SUDAAN runs for 2004 NNHS

Tables 1 through 3 show the distribution of standard error ratios (C_Xs) for estimates about facility level characteristics from the 2004 NNHS. The distributions are, respectively, for facility count estimates, facility average estimates for variables not correlated with bed size, and aggregate estimates of variables correlated with bed size. The variables are those for which data are available in the data dictionary compiled from the 2004 NNHS facility questionnaire [4].

Both Tables 1 and 2 show results for variables not correlated with bed size. For Table 1, SUDAAN's PROC CROSSTAB procedure was used with a NEST statement appropriate for the design. For Table 2, SUDAAN's PROC DESCRIPT was used. For each facility category listed in the tables, the estimation weight was identified in the WEIGHT statement to calculate estimates and the corresponding standard errors for 880 and 26 variables in Tables 1 and 2, respectively. Then, changing the WEIGHT statement to use the variance weight, we did the same calculations and noted the number of observations contributing to each estimate. We deleted any estimate with fewer than 30 observations because we considered such estimates unreliable. For example, estimates for an unusual facility characteristic within, say, region might be based on only 15 facilities possessing that rare characteristic, whereas 100 facilities in a region might possess the more common facility characteristic. Otherwise ratios were calculated using the standard error produced in the first run with estimation weights in the numerator and the standard error produced in the second run with the variance weights in the denominator. Single precision (with 8 digits) was used in these calculations. We then used SAS's PROC UNIVARIATE procedure to get the mean, median, and percentiles for the ratios, as shown in the tables. Estimates used in the distribution for each facility category in Table 1 were for counts of facilities by attributes such as ownership, chain membership status, etc. Estimates used in the distributions for Table 2 were for averages such as average percent of RNs with highest completed degree of MS/MSN per facility.

Table 3 is analogous to Tables 1 and 2 but pertains to the variables correlated with bed size. SUDAAN's PROC DESCRIPT was used to calculate weighted estimates and their variances for each of 75 bed size correlated variables (for example, number of certified beds, number of licensed beds, number of residents, number of lift devices, etc.) for each facility category shown.

For variables from Tables 1 through 3, the ratios of standard errors are in the following ranges, respectively:

$0.975 \le S_{Xe} / S_{Xv} \le 1.042$, $0.896 \le S_{Xe} / S_{Xv} \le 1.030$, and $0.954 \le S_{Xe} / S_{Xv} \le 1.081$.

Thus, from Section 3 above, the interval defined by $\hat{X} \pm 2.187 S_{Xe}$ (where 2.187=1.96/0.896) is the actual 95% confidence interval when C_X = 0.896 (from Table 2). Hence, to assure true 95% confidence intervals or 5% significance levels in hypothesis tests, one should use α = 0.0286 instead of 0.05 when using S_{Xe} instead of S_{Xv}.

Except for rounding errors, the C_{Xr} values for regional facility categories in Tables 1 and 3 appear to be approximately equal the values of their respective calibration ratios (given in Section 2) as expected.

5. Conclusion

Using ratio adjusted estimation weights to calculate standard errors will cause some over or underestimates of variance, especially for statistics by region. To evaluate standard errors calculated with estimation weights, ratios were formed using the standard errors calculated with estimation weights in the numerator and the standard errors calculated with variance weights in the denominator. As expected for aggregate type variables within region, these ratios were approximately equal (differed by 0.03 or less) the region specific calibration ratios used to produce the estimation weights. Hence, for aggregate estimates for the Northeast or the West regions where the calibration ratios exceed 1.0, the standard errors based on estimation weights overstate the standard errors. For the Midwest or the South regions where calibration ratios are less than 1.0, the standard errors of aggregate estimates based on estimation weights may understate the standard errors. For estimated averages in any region and for facility categories that are not restricted to individual regions, no predictions can be made regarding whether use of estimation weights in standard error calculations will result in understatement or overstatement of standard errors. The maximum ratio of standard errors calculated using the estimation weight to the standard errors from the variance weight is 1.042 for facility variables not correlated with bed size and 1.081 for bed size related facility variables. Ratios greater than 1 are conservative, and all estimates were less than a 10% overstatement. The minimum ratio is 0.896 for both variables not correlated with bed size and variables correlated with bed size. This underestimates the standard error. To assure a minimum of 95% in confidence intervals or a minimum of 5% significance levels in hypothesis tests when using S_{Xe} in place of S_{Xv}, one should lower the confidence level to α =0.0286 instead of 0.05.

Reference

[1] Rohatgi VK (1995). An Introduction to Probability Theory and Mathematical Statistics. John Wiley and Sons.

[2] NCHS (2006). 2004 National Nursing Home Survey and National Nursing Assistant Survey. http://www.cdc.gov/nchs/data/nnhsd/2004NNHS_DesignCollectionEstimates_072706tags.pdf . (Last visited on Sept. 27, 2007.)

[3] Shah, BV, Barnwell, BG, and Bieler, GS (1996). SUDAAN User's Manual, Release 7.0. Research Triangle Park, NC: Research Triangle Institute.

[4] NCHS (2006). 2004 National Nursing Home Survey Facility Questionnaire. http://www.cdc.gov/nchs/data/nnhsd/2004NNHS_Facility_Questionnaire_072506tags.pdf (Last visited on Sept. 27, 2007.)

Table 1: Distribution for NNHS ratios of standard errors for facility level variables not correlated with bed size: estimated counts of facilities

Facility Categories	Number of Estimates[1,2]	Minimum	5%	25%	Median	75%	95%	Maximum
Private	880	0.976	0.994	0.998	0.999	0.999	1.007	1.030
Non-profit	777	0.977	0.995	1.001	1.002	1.002	1.007	1.032
Government	620	0.977	0.987	0.991	0.991	0.991	0.996	1.012
Medicare & Medicaid	925	0.976	0.995	0.999	1.001	1.003	1.007	1.032
Medicare certified	405	0.992	0.995	0.997	0.997	0.997	0.999	1.002
Medicaid certified	569	0.979	0.988	0.992	0.992	0.992	0.996	1.000
Northeast	675	1.021	1.027	1.030	1.031	1.031	1.033	1.042
Midwest	781	0.968	0.973	0.976	0.976	0.977	0.979	0.987
South	779	0.988	0.993	0.995	0.995	0.995	0.997	1.000
West	687	1.001	1.004	1.007	1.007	1.008	1.010	1.015
Metro	895	0.976	0.997	1.001	1.008	1.022	1.022	1.030
Non-metro	808	0.978	0.987	0.991	0.993	0.994	0.998	1.042
Chain operated	823	0.975	0.993	0.997	0.998	0.998	1.004	1.030
Independent	839	0.978	0.994	0.999	1.000	1.001	1.006	1.032

[1] The estimates distributed for each facility category in this table are the estimated counts of facilities that are possible from the 2004 NNHS. These estimates are by individual categories listed as response options in multiple choice questions (such as type of facility ownership) and by categories described in binomial (yes/no) questions (such as "Is the medical director certified in geriatrics?"). These facility count estimates include (among others) those by ownership, certification status, bed size (<50, 50-99, 100-199, and 200+ beds), region, metro status, chain membership status, types of bed units offered (such as Alzheimer's, AIDS/HIV, behavior unit, specific diseases, and children), types of contracted services, methods for charging residents, services provided, what and how services (such as dental or mental health) are made available, special programs offered (such as hospice care), end of life programs engaged in, education levels and certification type for key staff (such as administrators, medical directors, and directors of nursing), tenure in year categories at this and at all facilities for key staff, methods of providing medical service, uses (such as physician orders or medication orders/dispensing) made of electronic information systems, recreation programs, flu vaccination programs, nursing staff retention/recruitment strategies, types of facility personnel on staff, reasons nursing staff worked overtime shifts, types of nursing staff employee benefits offered, types of staff belonging to labor unions, and duties performed by voluntary workers.

[2] Estimates based on fewer than 30 respondents were omitted from the distributions.

Table 2: Distribution for NNHS ratios of standard errors for facility level variables not correlated with bed size: estimated averages.

Facility Categories	Number of Estimates[1,2]	Minimum	5%	25%	Median	75%	95%	Maximum
Private	26	0.986	0.995	1.001	1.005	1.009	1.014	1.029
Non-profit	26	0.900	0.909	1.000	0.996	1.006	1.010	1.019
Government	25	0.976	0.988	0.994	0.999	1.005	1.011	1.022
Medicare & Medicaid	26	0.899	0.994	1.001	1.002	1.009	1.016	1.018
Medicare certified	20	0.978	0.982	0.993	0.998	1.003	1.009	1.009
Medicaid certified	23	0.951	0.961	0.998	0.999	1.006	1.019	1.020
Northeast	25	0.982	0.991	0.997	1.000	1.003	1.013	1.018
Midwest	26	0.912	0.920	0.996	0.993	1.001	1.009	1.014
South	26	0.989	0.995	0.999	1.001	1.004	1.005	1.005
West	25	0.993	0.993	0.996	0.998	0.999	1.004	1.006
Metro	26	0.982	0.998	1.000	1.003	1.006	1.016	1.017
Non-metro	26	0.905	0.984	0.997	0.999	1.007	1.011	1.023
Chain operated	26	0.896	0.992	1.001	1.000	1.006	1.015	1.030
Independent	26	0.978	0.991	0.999	1.003	1.008	1.012	1.026

[1] The estimates distributed for each facility category in this table are averages for: rates charged by patient source of payment, length of time in months spent at this and at all facilities for key staff (such as administrators, medical directors, and directors of nursing), days per month and days per week worked at this facility by medical directors, hourly wages of entry-level staff by staff type (RN, LPN, CAN, Aides/Orderlies), percent of facility RNs by highest education degree (Associated Degree, Diploma, BS/BSN, MS/MSN, specialty certifications) completed, and percent of facility nursing staff employed 1 year or more by nurse type.

[2] Estimates based on fewer than 30 respondents were omitted from the distributions.

Table 3: Distribution for NNHS ratio of standard errors for bed size related facility level variables: estimated aggregates.

Facility Categories	Number of Estimates[1,2]	Minimum	5%	25%	Median	75%	95%	Maximum
All	75	0.991	0.995	1.001	1.006	1.011	1.019	1.043
Private	75	0.978	0.992	1.001	1.006	1.010	1.019	1.049
Non-profit	73	0.960	0.991	1.000	1.002	1.006	1.014	1.017
Government	71	0.981	0.993	0.998	1.005	1.008	1.030	1.075
Medicare & Medicaid	75	0.976	0.995	1.001	1.005	1.009	1.017	1.043
Medicare certified	33	0.996	1.001	1.010	1.022	1.030	1.056	1.060
Medicaid certified	64	0.954	0.971	0.996	0.998	1.001	1.016	1.070
Northeast	72	1.000	1.000	1.005	1.011	1.015	1.029	1.036
Midwest	73	0.957	0.963	0.981	0.984	0.990	0.998	1.000
South	75	0.997	0.999	0.999	1.000	1.000	1.000	1.000
West	71	1.000	1.018	1.043	1.047	1.056	1.068	1.081
Metro	75	0.956	0.994	1.002	1.006	1.011	1.020	1.047
Non-metro	73	0.971	0.988	0.998	1.000	1.003	1.010	1.024
Chain operated	75	0.971	0.994	1.000	1.005	1.008	1.019	1.053
Independent	74	0.984	0.996	1.001	1.005	1.009	1.016	1.030
Beds<50	72	0.953	0.990	1.003	1.009	1.014	1.027	1.041
50-99 Beds	73	0.978	0.996	1.001	1.010	1.016	1.036	1.057
100-199 Beds	74	0.967	0.998	1.002	1.006	1.008	1.021	1.041
200+ Beds	61	0.953	0.987	0.999	0.999	1.002	1.010	1.019

[1] The estimates distributed for each facility category in this table are for aggregate estimates. The estimates include (among others) the estimated aggregates for: beds by certification status, residents by primary payment source, 2003 discharges and admissions, lift devices, employees who worked in the last week by nurse type and by full/part time status, FTE nurses who worked in the last week by nurse type and by employee/contract status, overtime shifts worked in the last week by type of nurse, RNs devoted solely to bedside care by full/part time status, and nurses hired or left the facility in the prior 3 months by nurse type.

[2] Estimates based on fewer than 30 respondents were omitted from the distributions.