## Estimating ASEC Variances with Replicate Weights

## Part I: Instructions for Using the ASEC Public Use Replicate Weight File to Create ASEC Variance Estimates

## Introduction

The U.S. Census Bureau releases a public use data file for the Current Population Survey’s Annual Social and Economic Supplement (ASEC) and a public use replicate weight file each fall. This document provides the data user with instructions on how to create the replicate weight estimates and how to use these estimates to calculate variances. Background information on how the ASEC replicate weights are created can be found in Part II.

## File Creation

The file CPS\_ASEC\_ASCII\_REPWGT\_yyyy.DAT, found on the U.S. Census Bureau website at: http://thedataweb.rm.census.gov/ftp/cps\_ftp.html, contains the replicate weights and match keys required to merge the replicate weight file to the public use survey data file (yyyy is the data collection year). This is an ASCII file with a record length of 1,617 columns. The following table documents the location of each variable. The weights, PWWGT0 – PWWGT160, are ten digits with four implied decimals. The match keys are H\_SEQ and PPPOS -- H\_SEQ with length 5 and PPPOS with length 2. Note these record lengths are different for the 2014 CPS ASEC sample with the redesigned income questions than for other ASEC samples.

|  |  |  |
| --- | --- | --- |
| Variable Name | **Start Column** | **Finish Column** |
| PWWGT0 (Full Sample Weight) | 1 | 10 |
| PWWGT1 | 11 | 20 |
| PWWGT2 | 21 | 30 |
| PWWGT3 | 31 | 40 |
|  |  |  |
| PWWGT(n) | 10n+1 | 10n+9 |
|  |  |  |
| PWWGT160 | 1601 | 1610 |
| H\_SEQ | 1611 | 1615 |
| PPPOS | 1616 | 1617 |

This file and the public use survey data file both have the full sample weight. On this file the variable name is PWWGT0, but on the public use survey data file the variable name is MARSUPWT. The full weight on this file is given as a means of verifying that the files are properly merged to the public use survey data.

The file CPS\_ASEC\_ASCII\_REPWGT\_yyyy.SAS, also found on the U.S. Census Bureau website, can be used as documentation while creating the replicate use weight file. This file provides SAS code that can be modified to create the replicate weight file. The file name and location need to be modified to meet the needs of the data users system and data file location. It also documents the location of each replicate weight and the two matching keys.

The file CPS\_ASEC\_ASCII\_REPWGT\_yyyy.SAS also provides the sum of each replicate weight across all records. These totals can be used for verification purposes. Sum each replicate weight across all records and then compare the totals to the sum of weights in this file to verify that the replicate weight file is created correctly.

## Merging the ASEC Replicate Weight File with the Person File

Obtain: ASEC Person File

 ASEC Replicate Weight File

Merge using H\_SEQ and PPPOS. This is a simple one-to-one match.

## Merging the ASEC Replicate Weight File with the Household File

Obtain: ASEC Person File

 ASEC Household File

 ASEC Replicate Weight File

Create a Reference Person File from the ASEC Person File, by keeping only records from the ASEC Person File with A\_EXPRRP = 1 or 2.

Create a Household/Reference Person File merging the ASEC Household File where H-HHTYPE = 1 and the Reference Person File by H\_SEQ (on the household file) and PH\_SEQ (on the person file).

Create a Reference Person Replicate Weight File from the ASEC Replicate Weight File by keeping only records from the ASEC Replicate Weight File with PPPOS = 41[[1]](#footnote-1). Merge this Reference Person Replicate Weight File with the Household/Reference Person File in a one-to-one match using the variable H\_SEQ.

## Merging the ASEC Replicate Weight File to the Family File

Obtain: ASEC Family File

 ASEC Person File

 ASEC Replicate Weight File

Create two new variables on the ASEC Person File. Set FH\_SEQ equal to the variable PH\_SEQ, and set FFPOS equal to the variable PHF\_SEQ. They will be used as match keys to the ASEC Family File. (At this point, you may want to keep any demographic variables for the reference person of the family.)

Merge the Person File with the ASEC Replicate Weight File using the variables H\_SEQ and PPPOS. Keep only the records with A\_FAMTYP = 1, 3, or 4 and A\_FAMREL = 1 or if A\_FAMTYP = 2 or 5. Only the records for the family reference person are required for the family file.

Create the Family File by merging the ASEC Person/Replicate Weight File with the ASEC Family File by FH\_SEQ and FFPOS. This is a simple one-to-one match.

## Creating Replicate Estimates

Replicate estimates are created using each of the 160 weights independently to create 160 replicate estimates. For point estimates, multiply the replicate weights by the item of interest at the record level (either an indicator variable to determine the number of people with a characteristic or a variable that contains some value, say, person income) and tally the weighted values to create the 160 replicate estimates. Use these replicate estimates in the formula to calculate the total variance for the item of interest. For example, say that the item of interest is the number of males in poverty. Tally the weights for all the records with variable A\_SEX = 1 and PERLIS = 1 to create the 160 replicate estimates of the number of males in poverty. Then use these estimates in the formula to calculate the total variance for the number of males.

The ASEC replicate weighting process may result in negative weights for some cases. Measures are taken in the full weighting process to ensure that the full sample weights do not result in negative weights. The replicate weights should be used in creating variances only and should not be used to create independent estimates.

## Use of Replicate Estimates in Variance Calculations

Calculate variance estimates for ASEC estimates using:

 (1)

where**is the estimate of the statistic of interest, such as a point estimate, ratio of domain means, regression coefficient, or log-odds ratio, using the weight for the full sample and**are the replicate estimates of the same statistic using the replicate weights. See reference Judkins (1990) and U.S. Census Bureau (2006) Chapter 14.

# Example for Total Variance of Point Estimates

The total variance for a point estimate**can be calculated by plugging the replicate weight estimates and the point estimate into formula (1):

,

where**are the replicate estimates.

Example for Variance of Regression Coefficients

Variances for regression coefficients**can be calculated using formula (1) as well. Calculating the 160 replicate regression coefficients and using formula (1),

,

gives the variance estimate for the regression coefficient**.

## Direct Variances Versus Generalized Variance Functions

Variances calculated using the above formulas often times do not match the variance estimates that are achieved by using generalized variance functions (GVF). The GVF is a simple model that expresses the variance as a function of the survey estimate. The parameters of the model are estimated using direct replicated variances from several estimates that have similar characteristics. These models provide a relatively easy way to obtain an approximate standard error on numerous characteristics.

Replicate weights can be used to calculate variances directly from the data using the formulas provided above. These variance estimates are considered to be direct variance estimates and are subject to some variance themselves.

## Examples of Calculating Variances Using:

## SAS, SUDAAN, or WesVar

## SAS CODE

The following is example SAS code that can be used to calculate standard errors using the replicate weights.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\* The FIRST STEP is to flag the data records \*;

\* desired after creating the SAS data sets. \*;

\* This example flags persons age 16+ and whom are male in poverty. \*;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

data user.data1;

 merge ASEC\_DATA\_2010 ASEC\_REPWGT\_2010;

 by h\_seq pppos;

 if a\_age>15 and a\_sex = 1 and perlis = 1 then male15\_plus\_pov = 1;

 else male15\_plus\_pov = 0;

run;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\* The SECOND STEP of code sums the full sample and the \*;

\* 160 replicate weights and writes them out to a file. \*;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

proc means data=user.data1 sum noprint;

 where male15\_plus\_pov=1;

 var marsupwt fmwgt1-fmwgt160;

 output out=user.data2 sum=est rw1-rw160;

run;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

\* The THIRD STEP of code uses the estimates of the full \*;

\* sample and the 160 replicates to compute the estimated \*;

\* replicate variance(s) using the formula(s) for 160 \*;

\* replicates. \*;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

data user.data3 (keep=char est var se cv);

 set user.data2 end=eof;

 if \_n\_=1 then sdiffsq = 0;

 array repwts{161} est rw1-rw160;

 do I = 2 to 161;

 sdiffsq = sdiffsq + (repwts{i} - repwts{1})\*\*2;

 end;

 if eof then do;

 var = (4/160) \* sdiffsq;

 se = (var)\*\*.5;

 cv = se/est;

 length char $20;

 char = 'Males 16+ in Poverty';

 output;

 end;

run;

proc print data = user.data3;

 var char est var se cv;

run;

##### SUDAAN CODE

The following is an example of SUDAAN code that can be used to calculate standard errors using the replicate weights.

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\* When specifying the sample design in SUDAAN the following \*

\* design statements need to be used: \*

\* IDVAR variables \*

\* REPWGT variables / ADJFAY = 4 \*

\* and \*

\* WEIGHT variable \*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/;

PROC CROSSTAB DATA = ASEC\_DATA\_2010 REPDATA = ASEC\_REPWGT\_2010 DESIGN = BRR;

IDVAR h\_seq pppos;

WEIGHT marsupwt;

REPWGT fmwgt1-fmwgt160 / ADJFAY = 4;

SUBPOPN 16 <= a\_age & a\_sex = 1 & perlis = 1;

TABLES \_one\_;

WESVAR

##### Using WesVar to calculate the variances for ASEC requires you to set up the WesVar data set properly. This can be done in the data file creation window of WesVar. This document will not walk you through all the steps required to use WesVar to calculate standard errors, but will assist you in the data creation window. There are five steps in creating your WesVar data set.

##### At the DATA FILE CREATION window in WesVar, add the full weight MARSUPWT to the full sample field.

##### Add the replicate weights FMWGT1 – FMWGT160 to the replicates field.

##### At the METHOD sidebar box, click on the FAY radio button.

##### In the FAY\_K window, enter 0.5 as the FAY adjustment value.

##### Add the variables of interest to the variables field.

After creating the WesVar data set, you can proceed with your analysis. The output pages of your analysis will contain the standard errors.

**Estimating ASEC Variances with Replicate Weights**

##### Part II: Replicate Variance Estimates for the ASEC

## Introduction

The variance of any survey estimate based on a probability sample may be estimated by the method of replication. This method requires that the sample selection, the collection of data, and the estimation procedures be independently carried through (replicated) several times. The dispersion of the resulting estimates then can be used to measure the variance of the full sample (reference [2]).

However, we would not consider repeating any large survey, such as the Annual Social and Economic (ASEC) supplement, several times to obtain variance estimates. A practical alternative is to draw a set of random samples from the full sample using the same principles of selection. We reuse the full sample several times by applying different weighting factors to the sample units. We treat these full samples as if they were different random samples and apply the estimation procedures to these random samples. We refer to these random samples as replicates.

For the ASEC, we used a total of 160 replicates to calculate the ASEC variance estimates. For additional information on determining the number of replicates see [2]. During the weighting processing, all 160 replicates undergo the same weighting adjustments.

In the following section we describe the methodology used in forming the 160 ASEC replicates. The theoretical basis of the methodology we use is based on the family of “balanced half-sample” methods. Wolter (1985) discusses this methodology in reference [3] and Fay (1995) extended the theory in reference [4]. We use both the balanced half-sample and the extended methodology to produce the replicated weights used for the ASEC supplement.

#### The Replication Method Applied to ASEC

The ASEC replicate weights are created differently for the self-representing (SR) strata and the nonself-representing strata (NSR). We derive both sets of replicate weights from methods known as “balanced half-sample” methods. The SR weights are created using the successive difference replication [4] and the NSR weights are created using the modified half sample technique [4].

Replicates for the ASEC are formed through a five-step process:

1. The first step is the construction of a *k* × *k* Hadamard matrix, where *k* is the number of replicates that will be formed.
2. Next, each SR case is assigned two rows of the Hadamard matrix and each NSR case one row.
3. In the third step, each sample case uses the assigned rows from the Hadamard matrix to calculate its replicate factors.
4. In the fourth step, the replicate factors are multiplied by the full-sample weights to produce the replicate weights.
5. Finally, the full sample and each of the replicate samples go through the weighting process.

At the end of this section, an example is provided to reinforce the steps of the replication method used for ASEC replicate weights. This example uses a sample of five cases and will create four replicates for each sample case.

##### Step 1: Construct the Hadamard Matrix

As mentioned earlier, the first step in creating the replicate weights for ASEC is the construction of a Hadamard matrix. A Hadamard matrix **H** is a *k* × *k* matrix with all elements either equal to +1 or -1. Hadamard matrices are unique in that they satisfy , where **I** is the identity matrix of order *k*, **H**k is a *k* × *k* Hadamard matrix, and **H**kT is the transpose of the *k* × *k* Hadamard matrix. The order *k* is necessarily 1, 2, or 4*t*, where *t* is a positive integer. An example of a 2 × 2 Hadamard matrix is as follows:

  (1)

Note that:



The Hadamard matrix allows us to choose certain replicate samples so that we can get an unbiased estimate of the variance with significantly fewer calculations than other half-sample methods (reference [3]). For ASEC, since 160 replicates are used, we used a 160 × 160Hadamard matrix to form our replicate factors. Please see reference [5] for information on the construction of 160 × 160 Hadamard matrices.

##### Step 2: Assign Row Values

Assignment of the row values depends on whether the sample case is SR or NSR. As mentioned earlier, replicate weights are formed differently for SR and NSR sample. Each SR case in the full sample will use two rows of the Hadamard matrix and the NSR cases are assigned to one row.

##### Assignment of Row Values for SR Cases

Since the first row of most Hadamard matrices consists entirely of +1=s, it is not assigned to a sample case. Therefore, the assignment process for the SR cases begins with the assignment of Rows 2 and 3 of the Hadamard matrix to the first sample case. The remaining row assignments are set up to ensure that consecutive sample cases share one row of the Hadamard matrix. Following this algorithm, Rows 3 and 4 are assigned to the second sample case. This row assignment continues until you reach the *kth* row of the *k* × *k* Hadamard matrix. At this point, you skip over the first row and return to the second row for the next assignment. After assigning all the row numbers incrementing by one, continue assigning the row numbers starting from Row 2, but increase the increment interval to two. Using an increment of two, the assignment process will continue with Rows 2 and 4 for the next sample case, followed by Rows 4 and 6, Rows 6 and 8, and so on. Under an increment of two, cycle through the rows twice to pick up all the row numbers. After assigning all increments of two, assign the row numbers with an increment of three. Use three cycles while incrementing by three. Continue to increase the increment and number of cycles up to a maximum increment of ten and then start the assignments over with the increment of one (if the independent sample is large enough to make this necessary). This provides 1,590 unique row assignment pairs.

**Assignment of Row Values for NSR Cases**

The NSR sampled strata are combined into pseudo-strata within each state to form paired strata. Each pseudo-stratum is assigned to a row of the Hadamard matrix. Within the pseudo-strata , one of the NSR PSU is randomly assigned the replicate factor 1.5 and the other NSR PSU receives the factor of 0.5. These values are assigned based on the Hadamard matrix. When the value of the Hadamard matrix changes the assigned replicate factor changes. For example, if the value of the Hadamard matrix is 1 and the first NSR PSU receives the replicate value of 1.5, the other NSR PSU receives a replicate factor of 0.5. When the value from the Hadamard matrix is –1, the first NSR PSU receives a replicate value of 0.5 and the second NSR PSU receives a replicate value of 1.5. These values are further adjusted to account for the unequal sizes of the original strata within pseudo-stratum.

In most cases the pseudo-strata consist of a pair of strata except where an odd number of strata within the state requires that a triplet be formed. In this case two rows of the Hadamard are assigned to the pseudo-stratum resulting in replicate factors of about 0.5, 1.7, and 0.8; or 1.5, 0.3, and 1.2 for the three PSUs assuming roughly equal sizes of the original strata. These values are further adjusted to account for the unequal sizes of the original strata within pseudo-stratum.

At the completion of the row assignment, each sample case will have *k* replicate factors - one factor for each replicate sample.

### Step 3: Calculation of the Replicate Factors for ASEC

The unique assignment of the row values to the SR sample cases ensures that the replicate factors take on one of three values: 0.3, 1.0, or 1.7. The replicate factors are calculated using the following formula:

  (2)

where

**i** = The sample case (i = 1, 2, …, n)

**r** = The replicate (r = 1, 2, …, *k*)

**RI** = The first row value assigned to sample case i

**RII** = The second row value assigned to sample case i

**h(RI, r)** = The (RI, r)th cell of the Hadamard matrix

**h(RII, r)** = The (RII, r)th cell of the Hadamard matrix

**NOTE:** The Hadamard cell to use is determined by the assigned row values and the column number corresponding to the replicate number. For example, when calculating a replicate factor for replicate 4, use the following cells from the Hadamard matrix: (RI, 4) and (RII, 4).

##### Step 4: Calculation of the Replicate Weights for ASEC

Each case within a probability sample has a sample weight that reflects the inverse of its probability of selection (i.e., the base weight). The weight can be viewed as the number of population members this sample case represents. The fourth step in the replication method calculates the replicate weights for each replicate sample. The replicate weights are calculated using the following formula:

  (3)

where

 **i** = The sample case ( i = 1, 2, …, n)

 **r** = The replicate sample (r = 1, 2, …, *k*)

 **Replicate Factor ir** = The replicate factor for the rth replicate of sample case i

 **Basewt i** = The full-sample base weight of sample case i

##### Step 5: Perform the Weighting Process

The final step in the creating replicate weights for ASEC involves sending the full sample and each replicate sample through the weighting process. The weighting process could be a simple ratio adjustment or could involve the implementation of a succession of complex adjustments. After the weighting adjustments, we are able to calculate estimates of variance for any ASEC estimate.

The base weights of ASEC sample cases went through the following adjustments[[2]](#footnote-2):

1. CPS Special weight Adjustment derived from CPS subsampling probabilities
2. Basic CPS Noninterview Adjustment
3. ASEC Noniterview Adjustment
4. First-stage Ratio Adjustment to reduce variances due to the sampling of NSR PSUs
5. SCHIP Adjustment factor for the over sampling selected demographic groups added to the ASEC sample
6. ASEC second-stage Ratio Adjustment to reduce variances by controlling to independent estimates of the current population
7. Family Equalization[[3]](#footnote-3)

Armed Forces members interviewed in households living off post or living on post with their families are in the ASEC estimates as long as there is one civilian adult (15+) living in the same household. The replicate weights assigned to the armed forces member are the same replicate weights assigned to the civilian householder.

##### An Example of the Replication Method

In an attempt to reinforce the steps of the replication method used for ASEC, we will create replicate samples for a sample data set. Our sample data set consists of five cases, all from an SR PSU, and we will create four replicates for each sample case. The sample cases and their corresponding full-sample weights are as follows (please assume the cases in this example are ordered in a manner reflective of the sample design):

**Table 1. Sample Data for ASEC Replication Example**

| **Sample Case** | **Sample Weight** |
| --- | --- |
| Case #1 | 15.00 |
| Case #2 | 23.00 |
| Case #3 | 19.00 |
| Case #4 | 16.00 |
| Case #5 | 21.00 |

Since we plan to create four replicates, we will need to construct a 4 × 4 Hadamard matrix. An example of a 4 × 4 Hadamard matrix is as follows:

  (4)

Recall that each sample case is assigned two rows of the Hadamard matrix. This assignment of rows begins with the second row and allows consecutive sample cases to share a row. The row assignments for the five sample cases are as follows:

**Table 2. Assignment of Rows for Sample Data**

| **Sample Case** | **Sample Weight** | **Row I** | **Row II** |
| --- | --- | --- | --- |
| Case #1 | 15.00 | 2 | 3 |
| Case #2 | 23.00 | 3 | 4 |
| Case #3 | 19.00 | 4 | 2 |
| Case #4 | 16.00 | 2 | 4 |
| Case #5 | 21.00 | 4 | 3 |

If we apply the Row I and Row II values into formula (2) for case #1, we arrive at the following replicate factors:



Recall that we determine the Hadamard matrix cell to use by the row value assigned in Table 2 and the column number corresponding to the replicate number. If we apply the Row I and Row II values to formula (2) for the remaining cases, the replicate factors will be as shown below in Table 3.

**Table 3. Replicate Factors for Sample Data**

| **Sample Case** | **Sample Weight** | **Replicate Factors** |
| --- | --- | --- |
| **Replicate 1** | **Replicate 2** | **Replicate 3** | **Replicate****4** |
| Case #1 | 15.00 | 1.0 | 0.3 | 1.7 | 1.0 |
| Case #2 | 23.00 | 1.0 | 1.7 | 1.0 | 0.3 |
| Case #3 | 19.00 | 1.0 | 1.0 | 0.3 | 1.7 |
| Case #4 | 16.00 | 1.0 | 1.0 | 1.7 | 0.3 |
| Case #5 | 21.00 | 1.0 | 0.3 | 1.0 | 1.7 |

Now that we have calculated the replicate factors for each sample case, we are ready to calculate the replicate weights using formula (3).

**Table 4. Replicate Weights for Sample Data**

| **Sample Case** | **Sample Weight** | **Replicate Weights** |
| --- | --- | --- |
| **Replicate 1** | **Replicate 2** | **Replicate 3** | **Replicate****4** |
| Case #1 | 15.00 | 15.00 | 4.50 | 25.50 | 15.00 |
| Case #2 | 23.00 | 23.00 | 39.10 | 23.00 |  6.90 |
| Case #3 | 19.00 | 19.00 | 19.00 | 5.70 | 32.30 |
| Case #4 | 16.00 | 16.00 | 16.00 | 27.20 | 4.80 |
| Case #5 | 21.00 | 21.00 | 6.30 | 21.00 |  35.70 |
| **Sum of Weights** | **94.00** | **94.00** | **84.90** | **102.40** | **94.70** |

The last step in the creation of the replicate weights is the implementation of any weighting adjustments. In our example, we use a ratio adjustment to control our sample to the population total of 100.00.

Therefore, we have to calculate a separate ratio adjustment factor for the full sample and for each replicate sample. In this example, the ratio adjustment factor formula is as follows:

  (5)

where

 **i** = The sample case (i = 1, 2, …, 5)

 **r** = The replicate sample (r = 0, 1, …, 4)

 ***NOTE: Replicate 0 refers to the full sample***

 ***w*i** = The weight for sample case i (either the full-sample weight or a replicate weight)

 **RAFr** = The ratio adjustment factor for replicate sample r

Using formula (5), the ratio adjustment factors for the full sample and each replicate sample are calculated as follows:

* Full Sample RAF = (100.00 ÷ 94.00) = **1.0638**
* Replicate 1 RAF = (100.00 ÷ 94.00) = **1.0638**
* Replicate 2 RAF = (100.00 ÷ 84.90) = **1.1779**
* Replicate 3 RAF = (100.00 ÷ 102.40) = **0.9766**
* Replicate 4 RAF = (100.00 ÷ 94.70) = **1.0560**

To perform the ratio adjustment, multiply the full-sample and replicate weights by the corresponding ratio adjustment factor. The following table provides the ratio adjusted weights.

**Table 5. Ratio Adjusted Weights for Sample Data**

| **Sample Case** | **Full****Sample Weight** | **Replicate Weights** |
| --- | --- | --- |
| **Replicate 1** | **Replicate 2** | **Replicate 3** | **Replicate****4** |
| Case #1 | 15.96 | 15.96 | 5.30 | 24.90 | 15.84 |
| Case #2 | 24.47 | 24.47 | 46.06 | 22.46 | 7.29 |
| Case #3 | 20.21 | 20.21 | 22.38 | 5.57 | 34.11 |
| Case #4 | 17.02 | 17.02 | 18.85 | 26.56 | 5.07 |
| Case #5 | 22.34 | 22.34 | 7.42 | 20.51 | 37.70 |
| **Sum of Weights** | **100.00** | **100.00** | **100.00** | **100.00** | **100.00** |

Using these ratio adjusted weights, we are ready to calculate estimates of variance for survey estimates. The next section discusses the calculation of variance estimates using replicates from the ASEC replication method.

**Variance Estimation using the ASEC Replication Weights**

Once the ASEC replicates weights are formed, estimates of variance for any full‑sample ASEC survey estimate is calculated using the following formula:

  (6)

where

 **r** = The replicate sample (r = 1......k)

 **o =** The full sample

 **k** = The total number of replicate samples (k = 160)

 **yo** = The survey estimate using the full-sample weights

 **yr** = The survey estimate using the replicate weights from replicate r

This variance estimate is the product of a constant and the sum of squared differences between each replicate survey estimate and the full-sample survey estimate.

##### An Example of Replicate Variance Estimation

Using the example set forth in the previous section, we will illustrate variance estimation using the ASEC replicate weights. Recall that our sample consists of five sample cases and four replicates per sample case. The goal of this section is to estimate the total number of employed persons in our population and its corresponding estimate of variance.

Assume that our five sample cases had the responses shown below in Table 6 when asked if they were employed during the time of interview.

**Table 6. Variance Estimation Using Sample Data**

| **Sample Case** | **Employed?** | **Sample Weight** | **Replicate Weights** |
| --- | --- | --- | --- |
| **Replicate 1** | **Replicate 2** | **Replicate 3** | **Replicate****4** |
| Case #1 | YES | **15.96** | **15.96** | **5.30** | **24.90** | **15.84** |
| Case #2 | NO | 24.47 | 24.47 | 46.06 | 22.46 | 7.29 |
| Case #3 | **YES** | **20.21** | **20.21** | **22.38** | **5.57** | **34.11** |
| Case #4 | **YES** | **17.02** | **17.02** | **18.85** | **26.56** | **5.07** |
| Case #5 | NO | 22.34 | 22.34 | 7.42 | 20.51 | 37.70 |

To calculate the full-sample survey estimate of the total employed population, we would add the full-sample weights of the sample cases that responded AYES@ to the employed question. Therefore, the total employed survey estimate is calculated as follows:

* Full-Sample Employed Estimate = 15.96 + 20.21 + 17.02 = **53.19**

In order to calculate the variance estimate for this survey estimate, we must calculate the same survey estimate for each of the replicate samples. The replicate survey estimates are as follows:

* Replicate 1 Employed Estimate = 15.96 + 20.21 + 17.02 = **53.19**
* Replicate 2 Employed Estimate = 5.30 + 22.38 + 18.85 = **46.53**
* Replicate 3 Employed Estimate = 24.90 + 5.57 + 26.56 = **57.03**
* Replicate 4 Employed Estimate = 15.84 + 34.11 + 5.07 = **55.02**

Now we can use these survey estimates in formula (6) to calculate the variance estimate for the total employed population. The calculation of this variance estimate is as follows:



Thus Var(yo) = **62.4501.**

Therefore, the survey estimate for total employed population in our example is **53.19** people. This survey estimate has an estimated variance of **62.4501**, or a standard error of **7.90** people.

## References

[1] Judkins, D. (1990) “Fay’s Method for Variance Estimation,” Journal of Official Statistics, Vol. 6, No. 3, 1990, pp.223-239.

[2] U.S. Census Bureau, Current Population Survey: Design and Methodology, Technical Paper 66 (October 2006), http://www.census.gov/prod/2006pubs/tp-66.pdf

[3] Wolter, Kirk (1985), Introduction to Variance Estimation, New York: Springer‑Verlag New York Inc.

[4] Fay, Robert, and Train, George (1995), “Aspects of Survey and Model‑Based Postcensal Estimation of Income and Poverty Characteristics for States and Counties,” Proceedings of the Section on Government Statistics, American Statistical Association, Alexandria, VA, pp. 154‑159.

[5] Plackett, R.L. and Burman, J.P. (1946), “The Design of Optimal Multifactorial Experiments,” Biometrika, 33, pp. 305-325.

1. If merging the 2005 file, the following records are reference persons as well:

 PPPOS=42 H\_SEQ=42236

 PPPOS=42 H\_SEQ=42241

 PPPOS=42 H\_SEQ=42387

 PPPOS=43 H\_SEQ=54922 [↑](#footnote-ref-1)
2. Further discussion and details of the ASEC weights can be found in reference [2], the Design and Methodology: Current Population Survey, Technical Paper 66. [↑](#footnote-ref-2)
3. The family equalization of replicate weighting may cause the replicate weights to be less than zero for some cases. This is caused by a negative ratio adjustment that is applied to the ASEC weight. During the creation of the ratio adjustment, adjusted counts are subtracted from control totals for some populations, which may result with a negative ratio causing the negative weights. [↑](#footnote-ref-3)