S2 Appendix. Estimation and variance of the total cost and cost per dose using design and calibrated weights.

c.rodriguez@auckland.ac.nz, stclr@channing.harvard.edu

February 14, 2019
1 Estimation of total cost $T$

The scientific goal is the estimation of a total. This total is composed of totals at different levels. For example, for the Brazilian immunization study there are costs that are specific to the regions, some other costs are specific to the municipalities and some others are specific to the facilities. We introduce the following notation in order to develop the details for a three multi-stage design.

Let $y_{Frmf}$ be the value specific to health facility $f$, in municipality $m$, in region $r$.

Let $y_{Strm}$ represents the value specific to municipality $m$, in region $r$.

Let $y_{fr}$ represents the value specific to region $r$.

In addition, let $U_r$ be the set of all regions, $U_{M_r}$ the set of all municipalities in region $r$ and $U_{F_{mr}}$ be the set of all facilities in municipality $m$, region $r$. There are in total $N_r$ regions, $N_{M_r}$ municipalities in region $r$ and $N_{F_{mr}}$ facilities in municipality $m$ in region $r$. The total cost can be represented as

$$T = T_c + T_f + T_m + T_n = T_c + \sum_{r=1}^{N_R} \sum_{m=1}^{N_{M_r}} \sum_{f=1}^{N_{F_{mr}}} y_{Frmf} + \sum_{r=1}^{N_R} \sum_{m=1}^{N_{M_r}} y_{Strm} + \sum_{r=1}^{N_R} y_{fr}. \quad (1)$$

where $T_c$ represents the additional cost accrued centrally and it is assumed known.

1.1 The design

To estimate (1), a multi-stage simple sampling design is used. Let $s_r$ represent the sub sample of regions, $s_{M_r}$ represent the sub sample of municipalities in region $r$ and $s_{F_{mr}}$ represent the sub sample facilities in municipality $m$ in region $r$, respectively. The design is as follows. At first stage, a simple sample, $s_r$, of $n_r$ regions is selected from $U_r$ with replacement. At second stage, for each region in $s_r$, a sample of $n_{M_r}$, municipalities is selected from $U_{M_r}$. Finally, at third stage, within each municipality in $s_{M_r}$, a sample of $n_{F_{mr}}$, facilities is selected from $U_{F_{mr}}$.

Each facility has a probability of being included in the final sample. This design can be characterized by its sampling probabilities. These probabilities are used in the estimation of the total $T$. This sampling probability for facility $f$ from municipality $m$, from region $r$, can be written as

$$\pi_{Frmf} = P(r \in s_r, m \in s_{M_r}, f \in s_{F_{mr}}) = P(f \in s_{F_{mr}} | r \in s_r, m \in s_{M_r})P(m \in s_{M_r} | r \in s_r)P(r \in s_r)$$

$$= \pi_f | m | r \pi_{fr} = \frac{n_{F_{mr}} n_{M_r} n_r}{N_{F_{mr}} N_{M_r} N_r}. \quad (2)$$

These probabilities can be written in this form only if we can assume that the sampling designs within each region was planned before the regions were selected and that the sampling design was implemented as planned, then we have a principle called invariance. Otherwise, it is not guaranteed that the probabilities can be calculated. Note that

$$I_{Frmf} = I_{Frmf} I_{Mrf} I_{fr}, \quad I_{Mrm} = I_{Mrm} I_{fr}, \quad I_{fr} = I_{fr}$$

where $I_{Frmf}, I_{Mrf}$ and $I_{fr}$ respectively represent the inclusion indicators for region $r$, municipality $m$ in region $r$ and facility $f$ in municipality $m$ in region $r$. The conditional inclusion probabilities are defined as $\pi_{f | mr} = E(\pi_{Frmf} | I_{Mrf}) = \frac{n_{F_{mr}}}{N_{F_{mr}}}$, $\pi_{rf} = E(I_{fr} | I_{fr}=1) = \frac{n_r}{N_r}$ and $\pi_{m | r} = E(I_{Mrm} | I_{fr}=1) = \frac{n_{M_r}}{N_{M_r}}$.

Using this, the number of facilities sampled in municipality, region $r$ can be written as $n_{F_{mr}} = \sum_{f=1}^{N_{F_{mr}}} I_{Frmf}$ and the number of facilities sampled in region $r$ can be written as $n_{Fr} = \sum_{m=1}^{N_{M_r}} I_{Mrm} \sum_{f=1}^{N_{F_{mr}}} I_{Frmf} I_{Frf}$.

1.2 Estimating $T_f$

We now present the estimation of the total and its standard error in the next sections.
• Under the design above, the weighted estimator of \( T_p \) is given by

\[
\hat{T}_p = \sum_{r=1}^{n_r} \sum_{m=1}^{n_{Mr}} \sum_{f=1}^{n_{Fmr}} w_{r,m,f} y_{m,f} = \sum_{r=1}^{n_r} \sum_{m=1}^{N_{Mr}} \sum_{f=1}^{N_{Fmr}} I_{r,m,f} w_{r,m,f} y_{r,m,f},
\]

where each \( w_{r,m,f} \) represents the sampling weight for facility \( f \) from municipality \( m \), from region \( r \).

• These weights are given by

\[
w_{r,m,f} = \frac{1}{\pi_{r,m,f}},
\]

where \( \pi_{r,m,f} \) is the inclusion probability for facility \( f \) in municipality \( m \) in region \( r \).

1.3 Estimating \( T_M \)

• Let \( w_{r,m} = \frac{1}{\pi_{m,r}} \); that is, the weight for municipality \( m \), which belongs to region \( r \).

• The weighted estimator of \( T_M \) is given by

\[
\hat{T}_M = \sum_{r=1}^{n_r} \sum_{m=1}^{n_{Mr}} w_{r,m} y_{r,m} = \sum_{r=1}^{n_r} \sum_{m=1}^{N_{Mr}} \sum_{f=1}^{N_{Fmr}} w_{r,m,f} w_{r,m,f} y_{r,m} = \sum_{r=1}^{n_r} \sum_{m=1}^{N_{Mr}} \sum_{f=1}^{N_{Fmr}} I_{r,m,f} w_{r,m,f} y_{r,m}.
\]

This estimator is unbiased for \( T_M \). The reason why we express (5) in terms of three sums is because it is used to calculate the variance.

1.4 Estimating \( T_R \)

• Let \( w_r = \frac{1}{\pi_r} \), that is, the weight corresponding to region \( r \).

• The weighted estimator of \( T_R \) is given by

\[
\hat{T}_R = \sum_{r=1}^{n_r} w_r y_r = \sum_{r=1}^{N_r} \sum_{m=1}^{N_{Mr}} \sum_{f=1}^{N_{Fmr}} w_{r,m,f} w_{r,m,f} y_{r,m} = \sum_{r=1}^{N_r} \sum_{m=1}^{N_{Mr}} \sum_{f=1}^{N_{Fmr}} I_{r,m,f} w_{r,m,f} y_{r,m}.
\]

This estimator is unbiased for \( T_R \).

1.5 The estimator of the total

Usually, IPW estimators only involve data at the last stage of the design, such as \( \hat{T}_p \). In this case, the variance of the estimator is well known and straightforward to calculate following standard properties of multi-stage designs [Sarndal et al. 1992]. Our estimator of interest, however, involves information at every stage of the design. In order to find an expression for the variance of \( \hat{T} \), we define the variables:

\[
w_{r,m,f} = y_{r,m,f} + \frac{w_{r,m,f} y_{r,m,f}}{w_{r,m,f} N_{Fmr}} y_{r,m} + \frac{w_r}{w_{r,m,f} N_{Fmr}} y_r + \frac{T_c}{n},
\]

\[
u_{r,m,f} = \frac{y_{r,m,f} + 1}{N_{Fmr}} y_{r,m} + \frac{1}{N_{Fmr}} y_r + \frac{T_c}{N}.
\]
Then, the estimator $\hat{T}$ be written as

$$\hat{T} = \sum_{r=1}^{N_R} I_{r_k} \sum_{m=1}^{N_{M_r}} I_{m_{r_m}} \sum_{f=1}^{N_{F_{m_f}}} I_{F_{m_f},} w_{F_{m_f},} u_{r_{m_f}}$$

and the total can be expressed as

$$T = \sum_{r=1}^{N_R} N_{M_r} N_{F_{m_f}} w_u^{*} u_{r_{m_f}}.$$  \hspace{1cm} (10)

Note that $E(\hat{T} | u) = T$.

2 Variance of $\hat{T}$

The variance of $E(\hat{T})$ can be calculated as

$$\text{Var} \left( \hat{T} \right) = E \left[ \text{Var} \left( \hat{T} | u \right) \right] + \text{Var} \left[ E \left( \hat{T} | u \right) \right],$$

where the first term represents the variation due to the sampling design and can be written as

$$E \left[ \text{Var} \left( \hat{T} | u \right) \right] = E \left[ \sum_{r=1}^{N_R} \sum_{s=1}^{N_{M_r}} \Delta_{n_r} I_{R_r} \hat{i}_{n_r} + \sum_{r=1}^{N_R} w_{r}, V_r \right],$$

where $\Delta_{n_r} = \pi_{n_r} - \pi_{r}, \pi_{n_r} = P(I_{n_r} = 1 | I_{R_r} = 1)$ which represents the pairwise inclusion probabilities at the regional level, $\hat{i}_{n_r} = w_{n_r} \sum_{m=1}^{N_{M_r}} \sum_{f=1}^{N_{F_{m_f}}} u_{r_{m_f}}$. The term $V_r$ is given by $V_r = \text{Var} \left( \hat{i}_{n_r} | u \right)$, where $\hat{i}_{n_r} = \sum_{m=1}^{N_{M_r}} \sum_{f=1}^{N_{F_{m_f}}} I_{F_{m_f},} w_{F_{m_f},} \pi_{n_r} u_{r_{m_f}}$. The term $V_n$ is calculated in the same way as $\hat{i}_{n_r}$.

For a three multi-stage design, we can express (12) as

$$E \left[ \text{Var} \left( \hat{T} | u \right) \right] = E \left[ V_n(\hat{T}) + V_M(\hat{T}) + V_F(\hat{T}) \right].$$

This expression is found using conditional expectations properties, where $V_n$ represents the variation due to the sampling of regions, $V_M$ represents the variation due to the sampling of municipalities $V_F$ is the variation due to the sampling of health facilities (see, e.g. Sarndal et al. [1992]). Under the underlying design, they have the forms

$$V_n = \left( \frac{1}{n_n} - \frac{1}{N_R} \right) N^2 \sum_{r} I_{r_k},$$

$$V_M = \frac{N_R}{n_n} \sum_{r \in U_n} \frac{N_{M_r}^2}{n_{M_r}} \left( \frac{1}{n_{M_r}} - \frac{1}{N_{M_r}} \right) S^2_{M_{m_r}},$$

$$V_F = \frac{N_R}{n_n} \sum_{r \in U_n} \frac{N_{M_r} N_{F_{m_f}}}{n_{M_r} n_{F_{m_f}}} \left( \frac{1}{n_{F_{m_f}}} - \frac{1}{N_{F_{m_f}}} \right) n_{F_{m_f}} S^2_{u_{m_f}},$$

where

$$S^2_{M_r} = \frac{1}{N_{M_r} - 1} \sum_{m \in U_{M_r}} \left( t_{m/r} - \bar{t}_r \right)^2; \quad t_{m/r} = \sum_{f \in U_{F_{m_f}}} u_{r_{m_f}} \quad \bar{t}_r = \frac{1}{N_R} \sum_{r \in U_n} t_r,$$

$$S^2_{M_{m_r}} = \frac{1}{N_{M_r} - 1} \sum_{m \in U_{M_r}} \left( t_{m/r} - \bar{t}_r \right)^2; \quad t_{m/r} = \sum_{f \in U_{F_{m_f}}} u_{r_{m_f}} \quad \bar{t}_r = \frac{1}{N_{M_r}} \sum_{m \in U_{M_r}} t_{m/r},$$

and

$$S^2_{u_{m_f}} = \frac{1}{N_{F_{m_f}} - 1} \sum_{f \in U_{F_{m_f}}} \left( u_{r_{m_f}} - \bar{u}_{m_f} \right)^2; \quad \bar{u}_{m_f} = \frac{1}{N_{F_{m_f}}} \sum_{f \in U_{F_{m_f}}} u_{r_{m_f}}.$$

The second term of (11) represents the probabilistic variation that would have been obtained had we had complete data. This can be expressed as
Estimation of the variance of \( \hat{T} \)

In practice, only the final sample of facilities is known, as so \( u_{rmf} \) is only known for the sample. This implies that the variance (13) cannot be calculated. However, we can estimate it using different approaches. The first approach is to analytically find an unbiased estimator. A second approach is to use bootstrap resampling, which return a consistent estimator.

3.1 Analytic estimation of \( E \left[ \text{Var} \left( \hat{T} \mid u \right) \right] \)

Under the underlying design, an unbiased estimator of the term (12) can be obtained analytically. An unbiased estimator of this is given by

\[
\hat{V} \hat{T} = \hat{V}_n \hat{T} + \hat{V}_m \hat{T} + \hat{V}_r \hat{T}.
\]  

(21)

\[
\hat{V}_n = \left( \frac{1}{n_R} - \frac{1}{N_R} \right) N^2 R^2 n_R^2,
\]

(22)

\[
\hat{V}_m = \frac{N_R}{n_R} \sum_{r \in s_R} N^2 m_r \left( \frac{1}{n_m} - \frac{1}{N_m} \right) \hat{S}^2 m_r,
\]

(23)

\[
\hat{V}_r = \frac{N_R}{n_R} \sum_{r \in s_R} N_m \sum_{m \in s_m} \left[ N^2 m_r \left( \frac{1}{n_m} - \frac{1}{N_m} \right) n_F m_r \hat{S}^2 m_r \right],
\]

(24)

where

\[
\hat{S}^2 m_r = \frac{1}{n_R - 1} \sum_{r \in s_R} \left( \hat{t}_r - \bar{t} \right)^2; \quad \hat{t}_r = \frac{1}{n_m} \sum_{m \in s_m} N_m \sum_{f \in s_F m_r} \frac{N_F m_r u_{rmf}}{n_F m_r}; \quad \bar{t} = \frac{1}{n_R} \sum_{r \in s_R} \hat{t}_r,
\]

(25)

\[
\hat{S}^2 m_r = \frac{1}{n_m - 1} \sum_{m \in s_m} \left( \hat{t}_m - \bar{t}_m \right)^2; \quad \hat{t}_m = \frac{1}{n_F m_r} \sum_{f \in s_F m_r} N_F m_r u_{rmf}; \quad \bar{t}_m = \frac{1}{n_m} \sum_{m \in s_m} \hat{t}_m.
\]

(26)

and

\[
\hat{S}^2 u_{rm} = \frac{1}{n_F m_r - 1} \sum_{f \in s_F m_r} \left( u_{rmf} - \hat{u}_{rm} \right)^2; \quad \hat{u}_{rm} = \frac{1}{n_F m_r} \sum_{f \in s_F m_r} u_{rmf}.
\]

(27)

3.2 Analytic estimation of \( \text{Var} \left[ E \left( \hat{T} \mid u \right) \right] \)

Under the underlying design, an unbiased estimator of the term (20) can be obtained analytically. An consistent estimator of this is given by

\[
\hat{\text{Var}} \left[ E \left( \hat{T} \mid u \right) \right] = \frac{N_R}{n_R} \sum_{r \in s_R} \sum_{m \in s_m} I_{r_m} \sum_{f \in s_F m_r} \frac{I_{r m f}}{\pi_{F m_r}} u_{rmf}^2 - N \left( \frac{1}{N} \sum_{r \in s_R} \sum_{m \in s_m} I_{r m} \sum_{f \in s_F m_r} \frac{I_{r m f}}{\pi_{F m_r}} u_{rmf} \right)^2.
\]

(28)

Or an unbiased estimator is given by

\[
\text{Var} \left[ E \left( \hat{T} \mid u \right) \right] = \sum_{r \in s_R} \sum_{m \in s_m} I_{r m} \sum_{f \in s_F m_r} \frac{I_{r m f} \hat{S}^2 m_r u_{rm}^2}{n_F m_r S^2 u_{rm}} \frac{S^2 m_r}{n_F m_r S^2 u_{rm}}.
\]

(29)
4 Post-design methods: Calibrated weights

The main purpose of post-designs methods is to improve the estimates by reducing the variance, i.e., to produce more efficient estimators. Post-design methods include techniques such as post stratification, estimation of weights and calibration of weights. The latter is known to yield more efficient estimates by adjusting the weights using information readily-available at first phase. In Brazil, for example, the total number of doses administered is known and this can be use to adjust the weights. 

Breslow et al. [2009a] and Breslow et al. [2009b] proposed using calibration to adjust the sampling weights in case-cohort designs. Let \( v_T, v_R \) and \( v_M \) denote readily available variables at first phase and at the different levels. The former is a variable available for all units at the final stage, the second is available for all second stage units and the latter is available for all first-stage units, e.g., regions. The main goal is to adjust weights such that estimates for the totals \( \hat{T}_{TP}, \hat{T}_{VM} \) and \( \hat{T}_{VR} \) equal the actual totals \( T_{TP}, T_{VM} \) and \( T_{VR} \), but forcing the new weights \( \hat{w} \) to be as close as possible to the design weights. For example, consider IPW estimates \( \hat{T}_{TP} \) and \( \hat{T}_{VM} \) based on the same weights used to estimate \( T \) with \( \hat{T} \), i.e.

\[
\hat{T}_{TP} = \sum_{r=1}^{n_R} \sum_{m=1}^{n_{M,r}} \sum_{f=1}^{n_{F,m}} w_{F,m,f} v_{F,m,f}; \quad \hat{T}_{VM} = \sum_{r=1}^{n_R} \sum_{m=1}^{n_{M,r}} w_{M,m} v_{M,m}; \quad \hat{T}_{VR} = \sum_{r=1}^{n_R} w_{VR} v_{VR}.
\]

Since we know \( T_{TP}, T_{VM} \) and \( T_{VR} \) we can modify the weights to force

\[
\hat{T}_{TP} = \sum_{r=1}^{n_R} \sum_{m=1}^{n_{M,r}} \sum_{f=1}^{n_{F,m}} \hat{w}_{F,m,f} v_{F,m,f} = T_{TP}; \quad \hat{T}_{VM} = \sum_{r=1}^{n_R} \sum_{m=1}^{n_{M,r}} \hat{w}_{M,m} v_{M,m} = T_{VM}; \quad \hat{T}_{VR} = \sum_{r=1}^{n_R} \hat{w}_{VR} v_{VR} = T_{VR}.
\] (30)

Intuitively, the modified weights \( \hat{w}_{F,m,f}, \hat{w}_{M,m} \) contain information about the entire population of facilities and we can use the new weights to find a new estimate of \( T \).

\[
\hat{T}_{cal} = \hat{T}_{T,cal} + \hat{T}_{M,cal} + \hat{T}_{R,cal} + T_C = \sum_{r=1}^{n_R} \sum_{m=1}^{n_{M,r}} \sum_{f=1}^{n_{F,m}} \hat{w}_{F,m,f} y_{F,m,f} + \sum_{r=1}^{n_R} \sum_{m=1}^{n_{M,r}} \hat{w}_{M,m} y_{M,m} + \sum_{r=1}^{n_R} \hat{w}_{VR} y_{VR} + T_C.
\]

Mathematically, a distance function \( d(w, \hat{w}) \) measuring the distance from original weights \( w \) to new weights \( \hat{w} \) is minimized subject to the constraints (30). There are several options for choosing this distance. Substantial gains in accuracy depend on how correlated the calibration variables are with the quantity of interest [Deville and Sarndal 1992]. In order to find the asymptotic variance of \( \hat{T}_{cal} \), we follow [Deville and Sarndal, 1992]. Choosing one the distance functions suggested by them ensures that we can write

\[
N_r^{-1} \left( \hat{T}_{F,cal} - \hat{T}_{F,reg} \right) = O_p \left( n_r^{-1} \right); \quad N_m^{-1} \left( \hat{T}_{M,cal} - \hat{T}_{M,reg} \right) = O_p \left( n_m^{-1} \right); \quad N_r^{-1} \left( \hat{T}_{R,cal} - \hat{T}_{R,reg} \right) = O_p \left( n_r^{-1} \right),
\] (31)

where \( \hat{T}_{F,reg}, \hat{T}_{M,reg} \) and \( \hat{T}_{R,reg} \) are the so-called regression estimators shown in [Deville and Sarndal] 1992 given by

\[
\hat{T}_{F,reg} = \sum_{r=1}^{n_R} \sum_{m=1}^{n_{M,r}} \sum_{f=1}^{n_{F,m}} \hat{w}_{F,m,f} y_{F,m,f}; \quad \hat{T}_{M,reg} = \sum_{r=1}^{n_R} \sum_{m=1}^{n_{M,r}} \hat{w}_{M,m} y_{M,m}; \quad \hat{T}_{R,reg} = \sum_{r=1}^{n_R} \hat{w}_{VR} y_{VR}.
\] (32)
and

\[ \hat{w}_{\text{Frm},f} = w_{\text{Frm},f} \left(1 + v_{\text{Frm},f}^T \hat{\lambda}_R \right); \quad \hat{\lambda}_R = \hat{T}_R^{-1} \left(T_{\text{vp}} - \hat{T}_{\text{vp}} \right); \quad \hat{T}_R = \sum_{r=1}^{N_R} \sum_{m=1}^{N_M} \sum_{f=1}^{N_{\text{Frm},f}} \hat{w}_{\text{Frm},f} v_{\text{Frm},f} v_{\text{Frm},f}^T; \]

\[ \hat{w}_{\text{Mrm}} = w_{\text{Frm}} \left(1 + v_{\text{Mrm}}^T \hat{\lambda}_M \right); \quad \hat{\lambda}_M = \hat{T}_M^{-1} \left(T_{\text{vm}} - \hat{T}_{\text{vm}} \right); \quad \hat{T}_M = \sum_{r=1}^{N_R} \sum_{m=1}^{N_M} \hat{w}_{\text{Mrm}} v_{\text{Mrm}} v_{\text{Mrm}}^T; \]

\[ \hat{w}_{\text{R}}, = w_{\text{R},r} \left(1 + v_{\text{R},r}^T \hat{\lambda}_R \right); \quad \hat{\lambda}_R = \hat{T}_R^{-1} \left(T_{\text{vp}} - \hat{T}_{\text{vp}} \right); \quad \hat{T}_R = \sum_{r=1}^{N_R} \hat{w}_{\text{R},r} v_{\text{R},r} v_{\text{R},r}^T. \]

Using (31) we can write:

\[ N_p^{-1} \left( \hat{T}_{\text{cal}} \right) = O_p \left(n_p^{-1} \right) + N_p^{-1} \left( \hat{T}_{\text{F,reg}} \right); \]

\[ N_p^{-1} \left( \hat{T}_{\text{reg}} \right) = O_p \left(n_p^{-1} \right) + N_p^{-1} \left( \hat{T}_{\text{M,reg}} + \hat{T}_{\text{R,reg}} \right); \]

\[ N_{\text{R}}^{-1} \left( \hat{T}_{\text{cal}} \right) = O_p \left(n_{\text{R}}^{-1} \right) + N_{\text{R}}^{-1} \left( \hat{T}_{\text{F,reg}} + O_p \left(n_{\text{M}}^{-1} \right) + N_{\text{M}}^{-1} \left( \hat{T}_{\text{M,reg}} \right) \right) + N_{\text{R}}^{-1} \left( \hat{T}_{\text{R,reg}} \right). \]

### 4.0.1 Variance estimation

In order to estimate the variance in practice we use the expression

\[ \hat{T}_{\text{cal}} \approx \hat{T}_{\text{F,reg}} + \hat{T}_{\text{M,reg}} + \hat{T}_{\text{R,reg}} + T_c \]

where, using Taylor expansion as in [Sarndal et al., 1992], we obtain that the approximate variance is the same as the variance of

\[ \sum_{r=1}^{N_R} \sum_{m=1}^{N_M} \sum_{f=1}^{N_{\text{Frm},f}} w_{\text{Frm},f} E_{\text{Frm},f} + \sum_{r=1}^{N_R} \sum_{m=1}^{N_M} \sum_{f=1}^{N_{\text{Mrm}}} w_{\text{Mrm}} E_{\text{Mrm}} + \sum_{r=1}^{N_R} w_{\text{R},r} E_{\text{R},r}, \]

with

\[ E_{\text{Frm},f} = y_{\text{Frm},f} - v_{\text{Frm},f}^T B_f; \quad B_f = T_f^{-1} \left( \sum_{r=1}^{N_R} \sum_{m=1}^{N_M} \sum_{f=1}^{N_{\text{Frm},f}} v_{\text{Frm},f} \right) \]

\[ E_{\text{Mrm}} = y_{\text{Mrm}} - v_{\text{Mrm}}^T B_f; \quad B_m = T_m^{-1} \left( \sum_{r=1}^{N_R} \sum_{m=1}^{N_M} v_{\text{Mrm}} \right) \]

\[ E_{\text{R},r} = y_{\text{R},r} - v_{\text{R},r}^T B_r; \quad B_r = T_r^{-1} \left( \sum_{r=1}^{N_R} v_{\text{R},r} \right). \]

We now define

\[ \hat{e}_{\text{Frm},f} = e_{\text{Frm},f} + \frac{w_{\text{Mrm}}}{w_{\text{Frm},f} n_{\text{Frm},f}} e_{\text{Mrm}} + \frac{w_{\text{R},r}}{w_{\text{Frm},f} n_{\text{Frm},f}} e_{\text{R},r}, \]

where \( e_{\text{Frm},f}, e_{\text{Mrm}} \) and \( e_{\text{R},r} \) are estimates for \( E_{\text{Frm},f}, E_{\text{Mrm}} \) and \( E_{\text{R},r} \). They are given by

\[ e_{\text{Frm},f} = y_{\text{Frm},f} - v_{\text{Frm},f}^T B_f; \quad B_f = T_f^{-1} \left( \sum_{r=1}^{N_R} \sum_{m=1}^{N_M} \sum_{f=1}^{N_{\text{Frm},f}} v_{\text{Frm},f} \right) \]

\[ e_{\text{Mrm}} = y_{\text{Mrm}} - v_{\text{Mrm}}^T B_f; \quad B_m = T_m^{-1} \left( \sum_{r=1}^{N_R} \sum_{m=1}^{N_M} v_{\text{Mrm}} \right) \]

\[ e_{\text{R},r} = y_{\text{R},r} - v_{\text{R},r}^T B_r; \quad B_r = T_r^{-1} \left( \sum_{r=1}^{N_R} v_{\text{R},r} \right). \]
Therefore

\[
\widetilde{\text{Var}} \left( \hat{T}_{\text{cal}} \right) = \text{Var} \left( \sum_{r=1}^{N_R} I_{R_r} \sum_{m=1}^{N_{M_r}} I_{M_{rm}} \sum_{f=1}^{N_{F_{rm}}} I_{F_{rmf}} w_{F_{rmf}} \tilde{u}_{rmf} \right) = \hat{E} \left[ \text{Var} \left( \hat{T}_e \mid u \right) \right] + \hat{E} \left[ \text{Var} \left( E \left( \hat{T}_e \mid u \right) \right) \right], \tag{40}
\]

with \( \hat{T}_e = \sum_{r=1}^{N_R} I_{R_r} \sum_{m=1}^{N_{M_r}} I_{M_{rm}} \sum_{f=1}^{N_{F_{rm}}} I_{F_{rmf}} w_{F_{rmf}} \tilde{u}_{rmf} \). Expression (40) is found in a similar way as the variance in section \( \text{2} \).
REFERENCES

N. Breslow, T. Lumley, C. Ballantyne, L. Chambless, and M Kulich. Improved horvitz-thompson estimation of model parameters from two-phase stratified samples: Applications in epidemiology. *Statistics in Biosciences*, 1(1):32–49, 2009a.

Norman E. Breslow, Thomas S. Lumley, Christie M. Ballantyne, Lloyd Ellwood Chambless, and Michal Kulich. Using the whole cohort in the analysis of case-cohort data. *American journal of epidemiology*, 169(11):1398–405, 2009b.

J.C. Deville and C. E. Sarndal. Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87:376–382, 1992.

C. E. Sarndal, B. Swensson, and J. Wretman. *Model Assisted Survey Sampling*. New York, 1992.