Multi-stage Sampling

1 Variance estimation for multi-stage sampling

Let U be a finite population and $\mathcal{P}(U)$ the power set of U. Let $p : \mathcal{S} \longrightarrow (0,1]$ be a r-stage sampling design ($r \geq 2$), where $\mathcal{S} \subset \mathcal{P}(U)$ is the set of all admissible samples under the design p. We express the hierarchical structure of the population U, with respect to the r-stage design p, as follows:

$$U \;=\; \bigsqcup_{i\in U^{(1)}} U_i^{(2+)} \;=\; \bigsqcup_{i\in U^{(1)}} \bigsqcup_{a\in U_i^{(2)}} U_{ia}^{(3)} \;=\; \cdots \;=\; \bigsqcup_{i\in U^{(1)}} \cdots \bigsqcup_{b\in U_{i\cdots}^{(r-1)}} U_{i\cdots b}^{(r)} \tag{1.1}$$

where $U^{(1)}$ is the set of all primary sampling units (PSU), and for each PSU $i \in U^{(1)}$, $U_i^{(2)}$ denotes the set of all secondary sampling units (SSU) contained in $i \in U^{(1)}$, and for each SSU $a \in U_i^{(2)}$, $U_{ia}^{(3)}$ denotes the set of all tertiary sampling units (TSU) contained in $a \in U_i^{(2)}$, and so on. Similarly, we express the hierachical structure of every admissible sample $s \in \mathcal{S}$ as follows:

$$s \;=\; \bigsqcup_{i\in s^{(1)}} s_i^{(2+)} \;=\; \bigsqcup_{i\in s^{(1)}} \bigsqcup_{a\in s_i^{(2)}} s_{ia}^{(3)} \;=\; \cdots \tag{1.2}$$

Let $y : U \longrightarrow \mathbb{R}$ be a population characteristic. Let T be its population total, i.e.

$$T \;=\; \sum_{u\in U} y_u \;=\; \sum_{i\in U^{(1)}} T_i \;=\; \sum_{i\in U^{(1)}} \sum_{a\in U_{ia}^{(2)}} T_{ia} \;=\; \cdots \;=\; \sum_{i\in U^{(1)}} \cdots \sum_{u\in U_{i\cdots}^{(r)}} y_u \tag{1.3}$$

Theorem 1.1
If $\widehat{T}_i : \mathcal{S}_i^{(2+)} \longrightarrow \mathbb{R}$ is an unbiased estimator for T_i, for each PSU *$i \in U^{(1)}$, then the random variable $\widehat{T} : \mathcal{S} \longrightarrow \mathbb{R}$ defined as follows:*

$$\widehat{T}(s) \;:=\; \sum_{i\in s^{(1)}} \frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}} \tag{1.4}$$

is a design-unbiased estimator for

$$T \;=\; \sum_{i\in U^{(1)}} T_i$$

If the r-stage sampling design has invariant and independent subsampling, then the design-variance of $\widehat{T}$ is given by:

$$\mathrm{Var}\Big[\,\widehat{T}\,\Big] \;=\; \underbrace{\mathrm{Var}^{(1)}\Big(E^{(2+)}\big(\,\widehat{T}\,\big|\,s^{(1)}\big)\Big)}_{\mathrm{V_{PSU}}} \;+\; \underbrace{E^{(1)}\Big(\mathrm{Var}^{(2+)}\big(\,\widehat{T}\,\big|\,s^{(1)}\big)\Big)}_{\mathrm{V_{subsampling}}}, \tag{1.5}$$

where

$$\begin{aligned}
\mathrm{Var}^{(1)}\Big(E^{(2+)}\big(\,\widehat{T}\,\big|\,s^{(1)}\big)\Big) \;&=\; \sum_{i\in U^{(1)}} \sum_{j\in U^{(1)}} \Delta_{ij}^{(1)} \frac{T_i}{\pi_i^{(1)}} \frac{T_j}{\pi_j^{(1)}}, \quad \text{and} \\
E^{(1)}\Big(\mathrm{Var}^{(2+)}\big(\,\widehat{T}\,\big|\,s^{(1)}\big)\Big) \;&=\; \sum_{i\in U^{(1)}} \frac{V_i}{\pi_i^{(1)}},
\end{aligned}$$

with

$$V_i \;:=\; \mathrm{Var}^{(2+)}\Big[\,\widehat{T}_i\,\Big] \quad\text{and}\quad \Delta_{ij}^{(1)} \;:=\; \begin{cases} \pi_i^{(1)}\left(1-\pi_i^{(1)}\right), & \text{if } i=j \\[2ex] \pi_{ij}^{(1)} - \pi_i^{(1)}\pi_j^{(1)}, & \text{if } i\neq j \end{cases} \tag{1.6}$$

Furthermore, if $\widehat{V}_i : \mathcal{S}_i^{(2+)} \longrightarrow \mathbb{R}$ *is an unbiased estimator for* $V_i := \mathrm{Var}^{(2+)}\left[\,\widehat{T}_i\,\right]$*, and* $\pi_i^{(1)} > 0$*,* $\pi_{ij}^{(1)} > 0$ *for any* PSU*s* $i, j \in U^{(1)}$*, then*

$$\begin{aligned}
\widehat{\mathrm{Var}}\left[\,\widehat{T}\,\right](s) &:= \sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}} \frac{\Delta_{ij}^{(1)}}{\pi_{ij}^{(1)}} \cdot \frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}} \cdot \frac{\widehat{T}_j(s_j^{(2+)})}{\pi_j^{(1)}} \;+\; \sum_{i\in s^{(1)}} \frac{\widehat{V}_i(s_i^{(2+)})}{\pi_i^{(1)}}, \\
\widehat{\mathrm{Var}}^{(1)}\left[\,\widehat{T}\,\right](s) &:= \sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}} \frac{\Delta_{ij}^{(1)}}{\pi_{ij}^{(1)}} \cdot \frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}} \cdot \frac{\widehat{T}_j(s_j^{(2+)})}{\pi_j^{(1)}} \;-\; \sum_{i\in s^{(1)}} \left(\frac{1}{\pi_i^{(1)}} - 1\right) \frac{\widehat{V}_i(s_i^{(2+)})}{\pi_i^{(1)}} \\
&= \sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}} \frac{\Delta_{ij}^{(1)}}{\pi_{ij}^{(1)}} \cdot \frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}} \cdot \frac{\widehat{T}_j(s_j^{(2+)})}{\pi_j^{(1)}} \;+\; \sum_{i\in s^{(1)}} \frac{\widehat{V}_i(s_i^{(2+)})}{\pi_i^{(1)}} \;-\; \sum_{i\in s^{(1)}} \frac{\widehat{V}_i(s_i^{(2+)})}{\left(\pi_i^{(1)}\right)^2} \\
\widehat{\mathrm{Var}}^{(2+)}\left[\,\widehat{T}\,\right](s) &:= \sum_{i\in s^{(1)}} \frac{\widehat{V}_i(s_i^{(2+)})}{\left(\pi_i^{(1)}\right)^2}
\end{aligned}$$

are unbiased estimators for $\mathrm{Var}\left[\,\widehat{T}\,\right]$, $\mathrm{V}_{\mathrm{PSU}} := \mathrm{Var}^{(1)}\left(E^{(2+)}\left(\,\widehat{T}\,\middle|\,s^{(1)}\right)\right)$*, and* $\mathrm{V}_{\mathrm{subsampling}} := E^{(1)}\left(\mathrm{Var}^{(2+)}\left(\,\widehat{T}\,\middle|\,s^{(1)}\right)\right)$*, respectively.*

Corollary 1.2

$$\widehat{\mathrm{Var}}^{(1)}\left[\,\widehat{T}\,\right](s) \;=\; \widehat{\mathrm{Var}}\left[\,\widehat{T}\,\right](s) \;-\; \widehat{\mathrm{Var}}^{(2+)}\left[\,\widehat{T}\,\right](s) \tag{1.7}$$

Proof of Theorem 1.1

$$\begin{aligned}
\mathrm{Var}^{(1)}\left[E^{(2+)}\left[\,\widehat{T}\,\middle|\,s^{(1)}\,\right]\right] &= \mathrm{Var}^{(1)}\left[E^{(2+)}\left[\sum_{i\in s^{(1)}} \frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}}\,\middle|\,s^{(1)}\right]\right] \\
&= \mathrm{Var}^{(1)}\left[\sum_{i\in s^{(1)}} \frac{E^{(2+)}\left[\,\widehat{T}_i(s_i^{(2+)})\,\right]}{\pi_i^{(1)}}\right] \\
&= \mathrm{Var}^{(1)}\left[\sum_{i\in s^{(1)}} \frac{T_i}{\pi_i^{(1)}}\right] \\
&= \sum_{i\in U^{(1)}}\sum_{j\in U^{(1)}} \Delta_{ij}^{(1)}\,\frac{T_i}{\pi_i^{(1)}}\,\frac{T_j}{\pi_j^{(1)}}
\end{aligned}$$

$$\begin{aligned}
E^{(1)}\left[\mathrm{Var}^{(2+)}\left[\,\widehat{T}\;\middle|\;s^{(1)}\,\right]\right]
&= E^{(1)}\left[\mathrm{Var}^{(2+)}\left[\,\sum_{i\in s^{(1)}}\frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}}\;\middle|\;s^{(1)}\,\right]\right] \\
&= E^{(1)}\left[\,\sum_{i\in s^{(1)}}\mathrm{Var}^{(2+)}\left[\,\frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}}\,\right]\right] \\
&= E^{(1)}\left[\,\sum_{i\in s^{(1)}}\frac{\mathrm{Var}^{(2+)}\left[\widehat{T}_i(s_i^{(2+)})\,\right]}{\left(\pi_i^{(1)}\right)^2}\right] \\
&= E^{(1)}\left[\,\sum_{i\in s^{(1)}}\frac{V_i/\pi_i^{(1)}}{\pi_i^{(1)}}\right] \\
&= \sum_{i\in U^{(1)}}\frac{V_i}{\pi_i^{(1)}}
\end{aligned}$$

$$\begin{aligned}
E\left(\widehat{\mathrm{Var}}^{(2+)}\left(\widehat{T}\right)\right)
&= E\left(\sum_{i\in s^{(1)}}\frac{\widehat{V}_i(s_i^{(2+)})}{\left(\pi_i^{(1)}\right)^2}\right)
= E^{(1)}\left(E^{(2+)}\left(\sum_{i\in s^{(1)}}\frac{\widehat{V}_i(s_i^{(2+)})/\pi_i^{(1)}}{\pi_i^{(1)}}\;\middle|\;s^{(1)}\right)\right) \\
&= E^{(1)}\left(\sum_{i\in s^{(1)}}\frac{E^{(2+)}\left[\,\widehat{V}_i(s_i^{(2+)})\;\middle|\;s^{(1)}\,\right]\Big/\,\pi_i^{(1)}}{\pi_i^{(1)}}\right)
= E^{(1)}\left(\sum_{i\in s^{(1)}}\frac{V_i/\,\pi_i^{(1)}}{\pi_i^{(1)}}\right)
= \sum_{i\in U^{(1)}}\frac{V_i}{\pi_i^{(1)}} \\
&= E^{(1)}\left(\mathrm{Var}^{(2+)}\left(\,\widehat{T}\;\middle|\;s^{(1)}\right)\right) = \mathrm{V}_{\mathrm{PSU}}
\end{aligned}$$

Similarly,

$$\begin{aligned}
E\left(\sum_{i\in s^{(1)}}\frac{\widehat{V}_i(s_i^{(2+)})}{\pi_i^{(1)}}\right)
&= E^{(1)}\left(E^{(2+)}\left(\sum_{i\in s^{(1)}}\frac{\widehat{V}_i(s_i^{(2+)})}{\pi_i^{(1)}}\;\middle|\;s^{(1)}\right)\right)
= E^{(1)}\left(\sum_{i\in s^{(1)}}\frac{E^{(2+)}\left[\,\widehat{V}_i(s_i^{(2+)})\;\middle|\;s^{(1)}\,\right]}{\pi_i^{(1)}}\right) \\
&= E^{(1)}\left(\sum_{i\in s^{(1)}}\frac{V_i}{\pi_i^{(1)}}\right) = \sum_{i\in U^{(1)}} V_i
\end{aligned}$$

Next, observe that

$$\begin{aligned}
E\left[\,\sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}}\frac{\Delta_{ij}^{(1)}}{\pi_{ij}^{(1)}}\cdot\frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}}\cdot\frac{\widehat{T}_j(s_j^{(2+)})}{\pi_j^{(1)}}\,\right]
&= E^{(1)}\left(E^{(2+)}\left(\sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}}\frac{\Delta_{ij}^{(1)}}{\pi_{ij}^{(1)}}\cdot\frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}}\cdot\frac{\widehat{T}_j(s_j^{(2+)})}{\pi_j^{(1)}}\;\middle|\;s^{(1)}\right)\right) \\
&= E^{(1)}\left(\sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}}\frac{\Delta_{ij}^{(1)}}{\pi_{ij}^{(1)}}\cdot\frac{E^{(2+)}\left[\,\widehat{T}_i(s_i^{(2+)})\cdot\widehat{T}_j(s_j^{(2+)})\;\middle|\;s^{(1)}\,\right]}{\pi_i^{(1)}\,\pi_j^{(1)}}\right)
\end{aligned}$$

Now, observe (the key technical observation) that

$$E^{(2+)}\left[\,\widehat{T}_i(s_i^{(2+)})\cdot\widehat{T}_j(s_j^{(2+)})\;\middle|\;s^{(1)}\,\right] = E^{(2+)}\left[\,\widehat{T}_i(s_i^{(2+)})\cdot\widehat{T}_j(s_j^{(2+)})\right] = \begin{cases}\mathrm{Var}^{(2+)}\left(\widehat{T}_i\right)+E^{(2+)}\left(\widehat{T}_i\right)^2, & \text{if } i=j,\\ E^{(2+)}\left(\widehat{T}_i\right)\cdot E^{(2+)}\left(\widehat{T}_i\right), & \text{if } i\neq j\end{cases}$$

Hence,

$$
\begin{aligned}
E\left[\,\sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}}\frac{\Delta_{ij}^{(1)}}{\pi_{ij}^{(1)}}\cdot\frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}}\cdot\frac{\widehat{T}_j(s_j^{(2+)})}{\pi_j^{(1)}}\,\right]
&= E^{(1)}\left(\sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}}\frac{\Delta_{ij}^{(1)}}{\pi_{ij}^{(1)}}\cdot\frac{E^{(2+)}\left[\,\widehat{T}_i(s_i^{(2+)})\cdot\widehat{T}_j(s_j^{(2+)})\;\middle|\;s^{(1)}\,\right]}{\pi_i^{(1)}\,\pi_j^{(1)}}\right) \\
&= E^{(1)}\left(\sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}}\frac{\Delta_{ij}^{(1)}}{\pi_{ij}^{(1)}}\cdot\frac{T_i}{\pi_i^{(1)}}\cdot\frac{Tj}{\pi_j^{(1)}} + \sum_{i\in s^{(1)}}\frac{(1-\pi_i^{(1)})\,V_i\big/\pi_i^{(1)}}{\pi_i^{(1)}}\right) \\
&= \sum_{i\in U^{(1)}}\sum_{j\in U^{(1)}}\Delta_{ij}^{(1)}\cdot\frac{T_i}{\pi_i^{(1)}}\cdot\frac{Tj}{\pi_j^{(1)}} + \sum_{i\in U^{(1)}}\frac{(1-\pi_i^{(1)})}{\pi_i^{(1)}}\cdot V_i \\
&= \sum_{i\in U^{(1)}}\sum_{j\in U^{(1)}}\Delta_{ij}^{(1)}\cdot\frac{T_i}{\pi_i^{(1)}}\cdot\frac{Tj}{\pi_j^{(1)}} + \sum_{i\in U^{(1)}}\frac{V_i}{\pi_i^{(1)}} - \sum_{i\in U^{(1)}}V_i
\end{aligned}
$$

We may now establish that

$$
\begin{aligned}
E\left[\,\widehat{\mathrm{Var}}\left(\widehat{T}\right)\,\right]
&= E\left[\sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}}\frac{\Delta_{ij}^{(1)}}{\pi_{ij}^{(1)}}\cdot\frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}}\cdot\frac{\widehat{T}_j(s_j^{(2+)})}{\pi_j^{(1)}} + \sum_{i\in s^{(1)}}\frac{\widehat{V}_i(s_i^{(2+)})}{\pi_i^{(1)}}\right] \\
&= \left\{\sum_{i\in U^{(1)}}\sum_{j\in U^{(1)}}\Delta_{ij}^{(1)}\cdot\frac{T_i}{\pi_i^{(1)}}\cdot\frac{Tj}{\pi_j^{(1)}} + \sum_{i\in U^{(1)}}\frac{V_i}{\pi_i^{(1)}} - \sum_{i\in U^{(1)}}V_i\right\} + \left\{\sum_{i\in U^{(1)}}V_i\right\} \\
&= \sum_{i\in U^{(1)}}\sum_{j\in U^{(1)}}\Delta_{ij}^{(1)}\cdot\frac{T_i}{\pi_i^{(1)}}\cdot\frac{Tj}{\pi_j^{(1)}} + \sum_{i\in U^{(1)}}\frac{V_i}{\pi_i^{(1)}} \\
&= \mathrm{Var}\left(\widehat{T}\right)
\end{aligned}
$$

Lastly, note that

$$
\widehat{\mathrm{Var}}^{(1)}\left[\widehat{T}\right](s) \;=\; \widehat{\mathrm{Var}}\left[\widehat{T}\right](s) \;-\; \widehat{\mathrm{Var}}^{(2+)}\left[\widehat{T}\right](s)
$$

Hence,

$$
\begin{aligned}
E\left[\,\widehat{\mathrm{Var}}^{(1)}\left(\widehat{T}\right)\right]
&= E\left[\,\widehat{\mathrm{Var}}\left(\widehat{T}\right)\right] - E\left[\,\widehat{\mathrm{Var}}^{(2+)}\left(\widehat{T}\right)\right] \\
&= \mathrm{Var}\left[\,\widehat{T}\,\right] - E^{(1)}\left(\mathrm{Var}^{(2+)}(\,\widehat{T}\,|\,s^{(1)})\right) \\
&= \mathrm{Var}^{(1)}\left(E^{(2+)}(\,\widehat{T}\,|\,s^{(1)})\right) + E^{(1)}\left(\mathrm{Var}^{(2+)}(\,\widehat{T}\,|\,s^{(1)})\right) - E^{(1)}\left(\mathrm{Var}^{(2+)}(\,\widehat{T}\,|\,s^{(1)})\right) \\
&= \mathrm{Var}^{(1)}\left(E^{(2+)}(\,\widehat{T}\,|\,s^{(1)})\right) = \mathrm{V}_{\text{subsampling}}
\end{aligned}
$$

$\square$

2 Variance estimation for three-stage sampling

Let U be a finite population and $\mathcal{P}(U)$ the power set of U. Let $p : \mathcal{S} \longrightarrow (0,1]$ be a three-stage sampling design, where $\mathcal{S} \subset \mathcal{P}(U)$ is the set of all admissible samples under the design p. We express the three-stage structure of the population U, with respect to the three-stage design p, as follows:

$$U \;=\; \bigsqcup_{i\in U^{(1)}} U_i^{(2+)} \;=\; \bigsqcup_{i\in U^{(1)}} \bigsqcup_{a\in U_i^{(2)}} U_{ia}^{(3)}, \tag{2.1}$$

where $U^{(1)}$ is the set of all primary sampling units (PSU), and for each PSU $i \in U^{(1)}$, $U_i^{(2)}$ denotes the set of all secondary sampling units (SSU) contained in $i \in U^{(1)}$, and for each SSU $a \in U_i^{(2)}$, $U_{ia}^{(3)}$ denotes the set of all tertiary sampling units (TSU) contained in $a \in U_i^{(2)}$. Similarly, we express the three-stage structure of every admissible sample $s \in \mathcal{S}$ as follows:

$$s \;=\; \bigsqcup_{i\in s^{(1)}} s_i^{(2+)} \;=\; \bigsqcup_{i\in s^{(1)}} \bigsqcup_{a\in s_i^{(2)}} s_{ia}^{(3)} \tag{2.2}$$

Let $y : U \longrightarrow \mathbb{R}$ be a population characteristic. Let T be its population total, i.e.

$$T \;=\; \sum_{u\in U} y_u \;=\; \sum_{i\in U^{(1)}} T_i \;=\; \sum_{i\in U^{(1)}} \sum_{a\in U_{ia}^{(2)}} T_{ia} \;=\; \sum_{i\in U^{(1)}} \sum_{a\in U_i^{(2)}} \sum_{u\in U_{ia}^{(3)}} y_u \tag{2.3}$$

Theorem 2.1

For a three-stage sampling design with invariant and independent subsampling,

1. *The random variable $\widehat{T} : \mathcal{S} \longrightarrow \mathbb{R}$, defined as follows:*

$$\widehat{T}(s) \;:=\; \sum_{i\in s^{(1)}} \frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}} \;=\; \sum_{i\in s^{(1)}} \frac{1}{\pi_i^{(1)}} \left(\sum_{a\in s_i^{(2)}} \frac{\widehat{T}_{ia}(s_{ia}^{(3)})}{\pi_{i|a}^{(2)}} \right) \;=\; \sum_{i\in s^{(1)}} \frac{1}{\pi_i^{(1)}} \left(\sum_{a\in s_i^{(2)}} \frac{1}{\pi_{i|a}^{(2)}} \left(\sum_{u\in s_{ia}^{(3)}} \frac{y_u}{\pi_{ia|u}^{(3)}} \right) \right)$$

   is a design-unbiased estimator for T, i.e. $E\left[\,\widehat{T}\,\right] = T$.

2. *If the three-stage sampling design has invariant and independent subsampling, then the design-variance of $\widehat{T}$ can be given by:*

$$\begin{aligned}
&\operatorname{Var}\left[\,\widehat{T}\,\right] \\
=\;& \operatorname{Var}^{(1)}\left(E^{(2+)}\left(\,\widehat{T}\;\middle|\; s^{(1)}\right)\right) \;+\; E^{(1)}\left(\operatorname{Var}^{(2+)}\left(\,\widehat{T}\;\middle|\; s^{(1)}\right)\right) \\
=\;& \operatorname{Var}^{(1)}\left(E^{(2+)}\left(\,\widehat{T}\;\middle|\; s^{(1)}\right)\right) \;+\; E^{(1)}\left\{\operatorname{Var}^{(2)}\left(E^{(3)}\left(\widehat{T}\;\middle|\; s^{(2)}\right)\;\middle|\; s^{(1)}\right) + E^{(2)}\left(\operatorname{Var}^{(3)}\left(\widehat{T}\;\middle|\; s^{(2)}\right)\;\middle|\; s^{(1)}\right)\right\} \\
=\;& \underbrace{\operatorname{Var}^{(1)}\left(E^{(2+)}\left(\,\widehat{T}\;\middle|\; s^{(1)}\right)\right)}_{\mathrm{V_{PSU}}} \;+\; \underbrace{E^{(1)}\left\{\operatorname{Var}^{(2)}\left(E^{(3)}\left(\widehat{T}\;\middle|\; s^{(2)}\right)\;\middle|\; s^{(1)}\right)\right\}}_{\mathrm{V_{SSU}}} + \underbrace{E^{(1)}\left\{E^{(2)}\left(\operatorname{Var}^{(3)}\left(\widehat{T}\;\middle|\; s^{(2)}\right)\;\middle|\; s^{(1)}\right)\right\}}_{\mathrm{V_{TSU}}}
\end{aligned}$$

   where

$$\mathrm{V_{PSU}} \;:=\; \operatorname{Var}^{(1)}\left(E^{(2+)}\left(\,\widehat{T}\;\middle|\; s^{(1)}\right)\right) \;=\; \sum_{i\in U^{(1)}} \sum_{j\in U^{(1)}} \Delta_{ij}^{(1)} \frac{T_i}{\pi_i^{(1)}} \frac{T_j}{\pi_j^{(1)}},$$

$$\mathrm{V_{SSU}} \;:=\; E^{(1)}\left\{\operatorname{Var}^{(2)}\left(\,E^{(3)}(\,\widehat{T}\;|\; s^{(2)}\,)\;\middle|\; s^{(1)}\right)\right\} \;=\; \sum_{i\in U^{(1)}} \frac{V_i^{(2)}}{\pi_i^{(1)}}, \quad \text{and}$$

$$\mathrm{V_{TSU}} \;:=\; E^{(1)}\left\{E^{(2)}\left(\,\operatorname{Var}^{(3)}(\,\widehat{T}\;|\; s^{(2)}\,)\;\middle|\; s^{(1)}\right)\right\} \;=\; \sum_{i\in U^{(1)}} \frac{1}{\pi_i^{(1)}} \left(\sum_{a\in U_i^{(2)}} \frac{V_{ia}^{(3)}}{\pi_{i|a}^{(2)}} \right),$$

with

$$\Delta_{ij}^{(1)} \;:=\; \left\{\begin{array}{ll} \pi_i^{(1)}\left(1-\pi_i^{(1)}\right), & \text{if } i=j \\ \\ \pi_{ij}^{(1)} - \pi_i^{(1)}\pi_j^{(1)}, & \text{if } i\neq j \end{array}\right.$$

$$V_i^{(2)} \;:=\; \mathrm{Var}^{(2)}\left[\,\sum_{a\in s_i^{(2)}} \frac{T_{ia}}{\pi_{i|a}^{(2)}}\,\right] \;=\; \sum_{a\in U_i^{(2)}}\sum_{b\in U_i^{(2)}} \Delta_{i|ab}^{(2)}\cdot\frac{T_{ia}}{\pi_{i|a}^{(2)}}\cdot\frac{T_{ib}}{\pi_{i|b}^{(2)}}$$

$$\Delta_{i|ab}^{(2)} \;:=\; \left\{\begin{array}{ll} \pi_{i|a}^{(2)}\left(1-\pi_{i|a}^{(2)}\right), & \text{if } a=b \\ \\ \pi_{i|ab}^{(2)} - \pi_{i|a}^{(2)}\,\pi_{i|b}^{(2)}, & \text{if } a\neq b \end{array}\right.$$

$$V_{ia}^{(3)} \;:=\; \mathrm{Var}^{(3)}\left(\,\widehat{T}_{ia}\;\middle|\;s^{(1)},s^{(2)}\,\right)$$

Theorem 2.2

For a three-stage sampling design $p:\mathcal{S}\subset\mathcal{P}(U)\longrightarrow\mathbb{R}$ with invariant and independent subsampling, let $\widehat{T}:\mathcal{S}\longrightarrow\mathbb{R}$ be the random variable defined as follows:

$$\widehat{T}(s) \;:=\; \sum_{i\in s^{(1)}}\frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}} \;=\; \sum_{i\in s^{(1)}}\frac{1}{\pi_i^{(1)}}\left(\sum_{a\in s_i^{(2)}}\frac{\widehat{T}_{ia}(s_{ia}^{(3)})}{\pi_{i|a}^{(2)}}\right) \;=\; \sum_{i\in s^{(1)}}\frac{1}{\pi_i^{(1)}}\left(\sum_{a\in s_i^{(2)}}\frac{1}{\pi_{i|a}^{(2)}}\left(\sum_{u\in s_{ia}^{(3)}}\frac{y_u}{\pi_{ia|u}^{(3)}}\right)\right)$$

Recall that $\widehat{T}$ is an unbiased estimator of the population total

$$T \;:=\; \sum_{i\in U^{(1)}} T_i \;:=\; \sum_{i\in U^{(1)}}\sum_{a\in U_{ia}^{(2)}} T_{ia} \;:=\; \sum_{i\in U^{(1)}}\sum_{a\in U_i^{(2)}}\sum_{u\in U_{ia}^{(3)}} y_u$$

Let $\widehat{\mathrm{Var}}\left[\,\widehat{T}\,\right]:\mathcal{S}\longrightarrow\mathbb{R}$ be the random variable defined in a recursive manner as follows:

$$\widehat{\mathrm{Var}}\left[\,\widehat{T}\,\right](s) \;:=\; \sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}}\frac{\Delta_{ij}^{(1)}}{\pi_{ij}^{(1)}}\cdot\frac{\widehat{T}_i(s_i^{(2+)})}{\pi_i^{(1)}}\cdot\frac{\widehat{T}_j(s_j^{(2+)})}{\pi_j^{(1)}} \;+\; \sum_{i\in s^{(1)}}\frac{\widehat{\mathrm{Var}}\left[\widehat{T}_i\right](s_i^{(2+)})}{\pi_i^{(1)}}$$

$$\widehat{\mathrm{Var}}\left[\widehat{T}_i\right](s_i^{(2+)}) \;:=\; \sum_{a\in s_i^{(2)}}\sum_{b\in s_i^{(2)}}\Delta_{i|ab}^{(2)}\cdot\frac{\widehat{T}_{ia}(s_{ia}^{(3)})}{\pi_{i|a}^{(2)}}\cdot\frac{\widehat{T}_{ib}(s_{ib}^{(3)})}{\pi_{i|b}^{(2)}} \;+\; \sum_{a\in s_i^{(2)}}\frac{\widehat{\mathrm{Var}}\left[\widehat{T}_{ia}\right](s_{ia}^{(3)})}{\pi_{i|a}^{(2)}}$$

$$\widehat{\mathrm{Var}}\left[\widehat{T}_{ia}\right](s_{ia}^{(3)}) \;:=\; \sum_{u\in s_{ia}^{(3)}}\sum_{v\in s_{ia}^{(3)}}\Delta_{ia|uv}^{(3)}\cdot\frac{y_u}{\pi_{ia|u}^{(3)}}\cdot\frac{y_v}{\pi_{ia|v}^{(3)}}$$

where

$$\widehat{T}_{ia}(s_{ia}^{(3)}) \;:=\; \sum_{u\in s_{ia}^{(3)}}\frac{y_u}{\pi_{ia|u}^{(3)}} \qquad\text{and}\qquad \widehat{T}_i(s_i^{(2+)}) \;:=\; \sum_{a\in s_i^{(2)}}\frac{\widehat{T}_{ia}(s_{ia}^{(3)})}{\pi_{i|a}^{(2)}} \;=\; \sum_{a\in s_i^{(2)}}\frac{1}{\pi_{i|a}^{(2)}}\left(\sum_{u\in s_{ia}^{(3)}}\frac{y_u}{\pi_{ia|u}^{(3)}}\right),$$

and

$$\Delta^{(1)}_{ij} \;:=\; \begin{cases} \pi^{(1)}_{i}\left(1-\pi^{(1)}_{i}\right), & \text{if } i=j \\ \\ \pi^{(1)}_{ij}-\pi^{(1)}_{i}\pi^{(1)}_{j}, & \text{if } i\neq j \end{cases}$$

$$\Delta^{(2)}_{i|ab} \;:=\; \begin{cases} \pi^{(2)}_{i|a}\left(1-\pi^{(2)}_{i|a}\right), & \text{if } a=b \\ \\ \pi^{(2)}_{i|ab}-\pi^{(2)}_{i|a}\,\pi^{(2)}_{i|b}, & \text{if } a\neq b \end{cases}$$

$$\Delta^{(3)}_{ia|uv} \;:=\; \begin{cases} \pi^{(3)}_{ia|u}\left(1-\pi^{(3)}_{ia|u}\right), & \text{if } u=v \\ \\ \pi^{(3)}_{ia|uv}-\pi^{(3)}_{ia|u}\,\pi^{(3)}_{ia|v}, & \text{if } u\neq v \end{cases}$$

Then, $\widehat{\mathrm{Var}}\left[\,\widehat{T}\,\right]$ *is a design-unbiased estimator of the design variance* $\mathrm{Var}\left[\,\widehat{T}\,\right]$ *of the* $\widehat{T}$*.*

Corollary 2.3

For a three-stage sampling design with invariant and independent subsampling, the fully expanded expression for $\widehat{\mathrm{Var}}\left[\,\widehat{T}\,\right]$ *is as follows:*

$$\begin{aligned}
\widehat{\mathrm{Var}}\left[\,\widehat{T}\,\right](s) \;&:=\; \sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}} \frac{\Delta^{(1)}_{ij}}{\pi^{(1)}_{ij}}\cdot\frac{\widehat{T}_i(s^{(2+)}_i)}{\pi^{(1)}_i}\cdot\frac{\widehat{T}_j(s^{(2+)}_j)}{\pi^{(1)}_j} \;+\; \sum_{i\in s^{(1)}}\frac{\widehat{V}_i(s^{(2+)}_i)}{\pi^{(1)}_i} \\
&= \sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}} \frac{\Delta^{(1)}_{ij}}{\pi^{(1)}_{ij}}\cdot\frac{\widehat{T}_i(s^{(2+)}_i)}{\pi^{(1)}_i}\cdot\frac{\widehat{T}_j(s^{(2+)}_j)}{\pi^{(1)}_j} \\
&\quad + \sum_{i\in s^{(1)}}\frac{1}{\pi^{(1)}_i}\left\{\sum_{a\in s^{(2)}_i}\sum_{b\in s^{(2)}_i}\Delta^{(2)}_{i|ab}\cdot\frac{\widehat{T}_{ia}(s^{(3)}_{ia})}{\pi^{(2)}_{i|a}}\cdot\frac{\widehat{T}_{ib}(s^{(3)}_{ib})}{\pi^{(2)}_{i|b}} + \sum_{a\in s^{(2)}_i}\frac{\widehat{V}_{ia}}{\pi^{(2)}_{i|a}}\right\} \\
&= \sum_{i\in s^{(1)}}\sum_{j\in s^{(1)}} \frac{\Delta^{(1)}_{ij}}{\pi^{(i)}_{ij}}\cdot\frac{\widehat{T}_i(s^{(2+)}_i)}{\pi^{(1)}_i}\cdot\frac{\widehat{T}_j(s^{(2+)}_j)}{\pi^{(1)}_j} \\
&\quad + \sum_{i\in s^{(1)}}\frac{1}{\pi^{(1)}_i}\left\{\sum_{a\in s^{(2)}_i}\sum_{b\in s^{(2)}_i}\Delta^{(2)}_{i|ab}\cdot\frac{\widehat{T}_{ia}(s^{(3)}_{ia})}{\pi^{(2)}_{i|a}}\cdot\frac{\widehat{T}_{ib}(s^{(3)}_{ib})}{\pi^{(2)}_{i|b}} + \sum_{a\in s^{(2)}_i}\frac{1}{\pi^{(2)}_{i|a}}\left(\sum_{u\in s^{(3)}_{ia}}\sum_{v\in s^{(3)}_{ia}}\Delta^{(3)}_{ia|uv}\frac{y_u}{\pi^{(3)}_{ia|u}}\frac{y_v}{\pi^{(3)}_{ia|v}}\right)\right\}
\end{aligned}$$

3 Variance estimation for three-stage sampling, with SRSWOR at each stage

First, recall that for a simple random sampling without replacement (SRSWOR), with fixed sample size n from a population of size N, the first- and second-order selection probabilities are given by:

$$\pi_i \;=\; \frac{n}{N} \quad \text{and} \quad \pi_{ij} \;=\; \frac{n(n-1)}{N(N-1)},$$

for any distinct units i, j in the population. The Horvitz-Thompson estimator of the population total of a population characteristic y is, by definition:

$$\widehat{T}_y^{\mathrm{HT}}(s) \;:=\; \frac{N}{n}\sum_{k\in s} y_k \;=\; w\cdot\sum_{k\in s} y_k, \quad \text{where } w \;:=\; \frac{N}{n}.$$

The design variance of $\widehat{T}_y^{\mathrm{HT}}$ is given by:

$$\begin{aligned}
\mathrm{Var}\Big[\,\widehat{T}_y^{\mathrm{HT}}\,\Big] &= N^2\left(1-\frac{n}{N}\right)\cdot\frac{1}{n}\left\{\frac{1}{N-1}\sum_{k\in U}(y_k-\overline{y}_U)^2\right\} \;=\; N^2\left(1-\frac{n}{N}\right)\cdot\frac{1}{n}\cdot\mathrm{SVar}\big(\{y_k\}_{k\in U}\big) \\
&= (nw)^2\left(1-\frac{1}{w}\right)\cdot\frac{1}{n}\cdot\mathrm{SVar}\big(\{y_k\}_{k\in U}\big) \;=\; nw^2\left(1-\frac{1}{w}\right)\cdot\mathrm{SVar}\big(\{y_k\}_{k\in U}\big) \\
&= nw\,(w-1)\cdot\mathrm{SVar}\big(\{y_k\}_{k\in U}\big)\,,
\end{aligned}$$

where $\overline{y}_U := \dfrac{1}{N}\sum_{k\in U} y_k$. Recall also that a design-unbiased estimator of $\widehat{T}_y^{\mathrm{HT}}$ is given by:

$$\begin{aligned}
\widehat{\mathrm{Var}}\Big[\,\widehat{T}_y^{\mathrm{HT}}\,\Big] &= N^2\left(1-\frac{n}{N}\right)\cdot\frac{1}{n}\left\{\frac{1}{n-1}\sum_{k\in s}(y_k-\overline{y}_s)^2\right\} \;=\; N^2\left(1-\frac{n}{N}\right)\cdot\frac{1}{n}\cdot\mathrm{SVar}\big(\{y_k\}_{k\in s}\big) \\
&= nw\,(w-1)\cdot\mathrm{SVar}\big(\{y_k\}_{k\in s}\big)
\end{aligned}$$

With the above observations, Corollary 2.3 immediately yields the following:

Corollary 3.1
For a three-stage sampling design with invariant and independent subsampling, where sampling random sampling without replacement (SRSWOR) is used at each stage, we have

$$\begin{aligned}
\widehat{\mathrm{Var}}\Big[\,\widehat{T}\,\Big](s) &= n^{(1)}w^{(1)}\Big(w^{(1)}-1\Big)\,\mathrm{SVar}\bigg(\Big\{\,\widehat{T}_i\,\Big\}_{i\in s^{(1)}}\bigg) \\
&\quad + w^{(1)}\sum_{i\in s^{(1)}}\underbrace{\left\{n_i^{(2)}w_i^{(2)}\Big(w_i^{(2)}-1\Big)\,\mathrm{SVar}\bigg(\Big\{\,\widehat{T}_{ia}\,\Big\}_{a\in s_i^{(2)}}\bigg) \;+\; w_i^{(2)}\sum_{a\in s_i^{(2)}} n_{ia}^{(3)}w_{ia}^{(3)}\Big(w_{ia}^{(3)}-1\Big)\,\mathrm{SVar}\Big(\{y_k\}_{k\in s_{ia}^{(3)}}\Big)\right\}}_{\widehat{V}_i^{(2+)}} \\
&= n^{(1)}w^{(1)}\Big(w^{(1)}-1\Big)\,\mathrm{SVar}\bigg(\Big\{\,\widehat{T}_i\,\Big\}_{i\in s^{(1)}}\bigg) \;+\; w^{(1)}\sum_{i\in s^{(1)}}\widehat{V}_i^{(2+)}
\end{aligned}$$

$$\begin{aligned}
\widehat{\mathrm{Var}}^{(1)}\left[\,\widehat{T}\,\right](s) &= n^{(1)}w^{(1)}\left(w^{(1)}-1\right)\mathrm{SVar}\left(\left\{\,\widehat{T}_i\,\right\}_{i\in s^{(1)}}\right) \;+\; w^{(1)}\sum_{i\in s^{(1)}}\widehat{V}_i^{(2+)} \;-\; \left(w^{(1)}\right)^2\sum_{i\in s^{(1)}}\widehat{V}_i^{(2+)} \\
&= n^{(1)}w^{(1)}\left(w^{(1)}-1\right)\mathrm{SVar}\left(\left\{\,\widehat{T}_i\,\right\}_{i\in s^{(1)}}\right) \;+\; w^{(1)}\left(1-w^{(1)}\right)\sum_{i\in s^{(1)}}\widehat{V}_i^{(2+)}
\end{aligned}$$

$$\widehat{\mathrm{Var}}^{(2+)}\left[\,\widehat{T}\,\right](s) = \left(w^{(1)}\right)^2\sum_{i\in s^{(1)}}\widehat{V}_i^{(2+)}$$