Some Proofs Associated with the Variance of the HT Estimator

The equality of the Sen-Yates-Grundy formula and the one we derived in class:

$$
\begin{aligned}
Var\left[\widetilde{a}_{HT}(\mathcal{S})\right] &= -\tfrac{1}{2}\textstyle\sum_{u\in\mathcal{P}}\sum_{v\in\mathcal{P}}\Delta_{uv}\left(\tfrac{y_u}{\pi_u}-\tfrac{y_v}{\pi_v}\right)^2\\
&= -\tfrac{1}{2}\textstyle\sum_{u\in\mathcal{P}}\sum_{v\in\mathcal{P}}\Delta_{uv}\left(\tfrac{y_u^2}{\pi_u^2}+\tfrac{y_v^2}{\pi_v^2}-\tfrac{2y_uy_v}{\pi_u\pi_v}\right)\\
&= -\tfrac{1}{2}\textstyle\sum_{u\in\mathcal{P}}\sum_{v\in\mathcal{P}}\Delta_{uv}\tfrac{y_u^2}{\pi_u^2}-\tfrac{1}{2}\sum_{u\in\mathcal{P}}\sum_{v\in\mathcal{P}}\Delta_{uv}\tfrac{y_v^2}{\pi_v^2}+\sum_{u\in\mathcal{P}}\sum_{v\in\mathcal{P}}\Delta_{uv}\tfrac{y_uy_v}{\pi_u\pi_v}\\
&= -\textstyle\sum_{u\in\mathcal{P}}\sum_{v\in\mathcal{P}}\Delta_{uv}\tfrac{y_u^2}{\pi_u^2}+\sum_{u\in\mathcal{P}}\sum_{v\in\mathcal{P}}\Delta_{uv}\tfrac{y_uy_v}{\pi_u\pi_v}
\end{aligned}
$$

Notice that the first term in the expression above can be re-written as:

$$
-\sum_{u\in\mathcal{P}}\sum_{v\in\mathcal{P}}\Delta_{uv}\frac{y_u^2}{\pi_u^2} = -\sum_{u\in\mathcal{P}}\left\{\frac{y_u^2}{\pi_u^2}\sum_{v\in\mathcal{P}}\Delta_{uv}\right\}
$$

It can also be shown that

$$
\begin{aligned}
\textstyle\sum_{v\in\mathcal{P}}\Delta_{uv} &= \textstyle\sum_{v\in\mathcal{P}}(\pi_{uv}-\pi_u\pi_v)\\
&= \textstyle\sum_{v\in\mathcal{P}}\pi_{uv}-\pi_u\sum_{v\in\mathcal{P}}\pi_v\\
&= n\pi_u-\pi_u n\\
&= 0
\end{aligned}
$$

Thus

$$
-\sum_{u\in\mathcal{P}}\sum_{v\in\mathcal{P}}\Delta_{uv}\frac{y_u^2}{\pi_u^2} = 0
$$

and so

$$
Var\left[\widetilde{a}_{HT}(\mathcal{S})\right] = -\frac{1}{2}\sum_{u\in\mathcal{P}}\sum_{v\in\mathcal{P}}\Delta_{uv}\left(\frac{y_u}{\pi_u}-\frac{y_v}{\pi_v}\right)^2 = \sum_{u\in\mathcal{P}}\sum_{v\in\mathcal{P}}\Delta_{uv}\frac{y_uy_v}{\pi_u\pi_v}
$$

Note that to verify $\sum_{v\in\mathcal{P}}\pi_v = n$ and $\sum_{v\in\mathcal{P}}\pi_{uv} = n\pi_u$ consider the following:

$$
\sum_{v\in\mathcal{P}}\pi_v = \sum_{v\in\mathcal{P}}E[D_v] = E\left[\sum_{v\in\mathcal{P}}D_v\right] = E[n] = n
$$

and

$$
\sum_{v\in\mathcal{P}}\pi_{uv} = \sum_{v\in\mathcal{P}}E[D_uD_v] = E\left[\sum_{v\in\mathcal{P}}D_uD_v\right] = E\left[D_u\sum_{v\in\mathcal{P}}D_v\right] = E[D_u\times n] = nE[D_u] = n\pi_u
$$

The variance of the HT estimator under SRSWOR

$$
\begin{aligned}
Var\left[\widetilde{a}_{HT}(\mathcal{S})\right] &= \sum_{u\in\mathcal{P}} \sum_{v\in\mathcal{P}} \Delta_{uv} \frac{y_u y_v}{\pi_u \pi_v} \\
&= \sum_{u\in\mathcal{P}} \Delta_{uu} \frac{y_u^2}{\pi_u^2} + \sum_{u\in\mathcal{P}} \sum_{v\neq u} \Delta_{uv} \frac{y_u y_v}{\pi_u \pi_v} \\
&= \sum_{u\in\mathcal{P}} \frac{y_u^2(1-\pi_u)}{\pi_u} + \sum_{u\in\mathcal{P}} \sum_{v\neq u} \left(\frac{\pi_{uv}}{\pi_u \pi_v} - 1\right) y_u y_v
\end{aligned}
$$

Substituting $\pi_u = \frac{n}{N}$ and $\pi_{uv} = \frac{n(n-1)}{N(N-1)}$ yields the following:

$$
\begin{aligned}
Var\left[\widetilde{a}_{HT}(\mathcal{S})\right] &= \sum_{u\in\mathcal{P}} y_u^2 \left(\frac{N-n}{n}\right) + \sum_{u\in\mathcal{P}} \sum_{v\neq u} \left(\frac{N(n-1)}{n(N-1)} - 1\right) y_u y_v \\
&= \sum_{u\in\mathcal{P}} y_u^2 \left(\frac{N-n}{n}\right) + \sum_{u\in\mathcal{P}} \sum_{v\neq u} \left(\frac{n-N}{n(N-1)}\right) y_u y_v \\
&= \left(\frac{N-n}{n}\right) \left[\sum_{u\in\mathcal{P}} y_u^2 - \frac{1}{N-1} \sum_{u\in\mathcal{P}} \sum_{v\neq u} y_u y_v\right] \\
&= \left(\frac{N-n}{n}\right) \left[\sum_{u\in\mathcal{P}} y_u^2 - \frac{1}{N-1} \sum_{u\in\mathcal{P}} \sum_{v\neq u} y_u y_v - \frac{1}{N-1} \sum_{u\in\mathcal{P}} y_u^2 + \frac{1}{N-1} \sum_{u\in\mathcal{P}} y_u^2\right] \\
&= \left(\frac{N-n}{n}\right) \left[\sum_{u\in\mathcal{P}} y_u^2 - \frac{1}{N-1} \sum_{u\in\mathcal{P}} \sum_{v\in \surd} y_u y_v + \frac{1}{N-1} \sum_{u\in\mathcal{P}} y_u^2\right] \\
&= \left(\frac{N-n}{n}\right) \left[\frac{N}{N-1} \sum_{u\in\mathcal{P}} y_u^2 - \frac{1}{N-1} \left(\sum_{u\in\mathcal{P}} y_u\right)^2\right] \\
&= \left(\frac{N-n}{N-1}\right) \left(\frac{N}{n}\right) \left[\sum_{u\in\mathcal{P}} y_u^2 - N\overline{y}^2\right] \\
&= \left(\frac{N-n}{N-1}\right) \left(\frac{N}{n}\right) \sum_{u\in\mathcal{P}} (y_u - \overline{y})^2
\end{aligned}
$$