BIOSTAT 880 HW3 Solution, Fall 2024

Yize Hao

Solution compiled on October 3, 2024

5.9

Consider a simple random sample of size n with r respondents and $m = n - r$ nonrespondents, and let $\bar{y}_R$ and s_R^2 be the sample mean and variance of the respondents' data, and $\bar{y}_{NR}$ and s_{NR}^2 be the sample mean and variance of the imputed data. Show that the mean and variance $\bar{y}_*$ and s_*^2 of all the data can be written as

$$\bar{y}_* = \frac{(\bar{y}_R + m\bar{y}_{NR})}{n}$$

and

$$s_*^2 = \frac{(r-1)s_R^2 + (m-1)s_{NR}^2 + rm(\bar{y}_R - \bar{y}_{NR})^2/n}{n-1}.$$

$$\bar{y}_R = \frac{\sum_{i=1}^r y_{iR}}{r} \quad \text{and} \quad \bar{y}_{NR} = \frac{\sum_{i=1}^m y_{iNR}}{m}$$

So,

$$\bar{y}_* = \frac{r}{n}\bar{y}_R + \frac{m}{n}\bar{y}_{NR}.$$

$$s_*^2 = \frac{1}{n-1}\left[\sum_{i=1}^r (y_{iR} - \bar{y}_*)^2 + \sum_{i=1}^m (y_{iNR} - \bar{y}_*)^2\right]$$

Expanding,

$$s_*^2 = \frac{1}{n-1}\left[\sum_{i=1}^r (y_{iR} - \bar{y}_R + \bar{y}_R - \bar{y}_*)^2 + \sum_{i=1}^m (y_{iNR} - \bar{y}_{NR} + \bar{y}_{NR} - \bar{y}_*)^2\right]$$

$$s_*^2 = \frac{1}{n-1}\left[\sum_{i=1}^r (y_{iR} - \bar{y}_R)^2 + \sum_{i=1}^m (y_{iNR} - \bar{y}_{NR})^2\right] + \frac{rm}{n}(\bar{y}_R - \bar{y}_{NR})^2$$

Using $s_R^2 = \frac{1}{r-1}\sum_{i=1}^r (y_{iR} - \bar{y}_R)^2$ and $s_{NR}^2 = \frac{1}{m-1}\sum_{i=1}^m (y_{iNR} - \bar{y}_{NR})^2$,

$$s_*^2 = \frac{(r-1)s_R^2 + (m-1)s_{NR}^2 + \frac{rm}{n}(\bar{y}_R - \bar{y}_{NR})^2}{n-1}$$

And,

$$\frac{rm}{n} \cdot \frac{1}{n-1} = \frac{rm}{n(n-1)}.$$

Therefore,

$$s_*^2 = \frac{(r-1)s_R^2 + (m-1)s_{NR}^2 + \frac{rm}{n}(\bar{y}_R - \bar{y}_{NR})^2}{n-1}.$$

5.10

(a)

$$\bar{y}_* = \frac{r}{n}\bar{y}_R + \frac{m}{n}\bar{y}_{NR}$$

So,

$$E[\bar{y}_*] = \frac{r}{n}E[\bar{y}_R] + \frac{m}{n}E[\bar{y}_{NR}]$$

$$E[\bar{y}_R] = E\left[\frac{1}{r}\sum_{i=1}^{r} y_{iR}\right]$$

$$E[\bar{y}_R] = \bar{Y}$$

So,

$$E[\bar{y}_*] = \frac{r}{n}\bar{Y}$$

Similarly,

$$E[\bar{y}_{NR}] = \bar{Y} \quad \text{(due to draw with replacement)}$$

So,

$$E[\bar{y}_*] = \bar{Y}.$$

(b)

$$\mathrm{Var}(y_{NR}|y_R) = \mathrm{Var}\left(\frac{r}{n}y_R + \frac{m}{n}y_{NR}|y_R\right)$$

$$= \frac{m^2}{n^2}\mathrm{Var}(y_{NR}|y_R)$$

$$= \frac{m}{n} \cdot \mathrm{Var}\left(\frac{1}{m}\sum_{i=1}^{m} y_{iNR}|y_R\right)$$

Where $y_{iNR} \overset{iid}{\sim} y_R$ with $P[y_{iNR} = y_{jR}] = \frac{1}{r}, j = 1, \ldots, r$.

$$\text{Var}(y_{iNR}) = E[y_{iNR}^2] - E[y_{iNR}]^2$$

$$E[y_{iNR}] = \frac{1}{r}\sum_{i=1}^{r} P(y_{iNR} = y_R) y_{iR}$$

$$= \frac{r}{n}\sum_{i=1}^{r} y_{iR}$$

$$= \bar{y}_R \quad \text{(as expected)}$$

$$P(y_{iNR} = \bar{y}_R) = \frac{r}{n}$$

So,

$$E[y_{iNR}^2] = \frac{1}{r}\sum_{i=1}^{r} y_{iR}^2$$

Thus,

$$\text{Var}(y_{iNR}) = E[y_{iNR}^2] - E[y_{iNR}]^2$$

$$= \frac{1}{r}\sum_{i=1}^{r}(y_{iR}^2 - \bar{y}_R^2)$$

$$= \frac{1}{r}\sum_{i=1}^{r}(y_{iR} - \bar{y}_R)^2$$

Therefore,

$$\text{Var}(y_{NR}|y_R) = \frac{1}{r}\sum_{i=1}^{r}(y_{iR} - \bar{y}_R)^2$$

Thus,

$$\text{Var}(y_{iR}|y_R) = \frac{(r-1)s_R^2}{r}$$

$$= \frac{1}{r}\sum_{i=1}^{r}(y_{iR} - \bar{y}_R)^2 + 2\sum_{i=1}^{r}(y_{iR} - \bar{y}_R)(\bar{y}_R - \bar{Y}_R)$$

$$= \frac{1}{r}\sum_{i=1}^{r}(y_{iR} - \bar{y}_R)^2 + 2\sum_{i=1}^{r}(y_{iR} - \bar{y}_R) \cdot 0$$

$$= \frac{(r-1)}{r} s_R^2$$

Thus,

$$\mathrm{Var}(y_{NR}|y_R) = \mathrm{Var}\left(\frac{r}{n}\bar{y}_R + \frac{m}{n}\bar{y}_{NR}|y_R\right)$$

$$= \frac{m^2}{n^2}\mathrm{Var}(y_{NR}|y_R)$$

$$= \frac{m}{n} \cdot \mathrm{Var}\left(\frac{1}{m}\sum_{i=1}^{m} y_{iNR}|y_R\right)$$

$$= \frac{m}{n} \cdot \mathrm{Var}\left(\frac{1}{m}\sum_{i=1}^{m} y_{iR}\right)$$

$$= \frac{m}{n} \cdot \frac{1}{m} s_R^2(1 - r^{-1})$$

Thus,

$$\boxed{\mathrm{Var}(y_{NR}|y_R) = \frac{m}{n} s_R^2(1 - r^{-1})}$$

Similarly:

$$E[S_*^2|y_R] = E\left(\frac{(r-1)s_R^2 + (m-1)s_{NR}^2 + \frac{rm}{n}(\bar{y}_R - \bar{y}_{NR})^2}{n-1}\middle| y_R\right)$$

$$= \frac{(r-1)}{n-1}s_R^2 + \frac{(m-1)}{n-1}E[s_{NR}^2|y_R] + \frac{rm}{n(n-1)}E[(\bar{y}_R - \bar{y}_{NR})^2|y_R]$$

$$E[s_{NR}^2|y_R] = E\left(\frac{1}{m-1}\sum_{j=1}^{m}(y_{jNR} - \bar{y}_{NR})^2|y_R\right)$$

$$= \frac{1}{m-1}\sum_{j=1}^{m} E[(y_{jNR} - \bar{y}_{NR})^2|y_R]$$

And,

$$E[(\bar{y}_R - \bar{y}_{NR})^2|y_R] = E[\bar{y}_R|y_R] - E[\bar{y}_{NR}|y_R]$$

$$= E[\bar{y}_R|y_R] - \frac{1}{r}\sum_{j=1}^{r} E[\bar{y}_{NR}|y_R]$$

$$\Rightarrow E[\bar{y}_R - \bar{y}_{NR}] = 0$$

And,

$$\mathrm{Var}(\bar{y}_R - \bar{y}_{NR}|y_R)$$

$$= \mathrm{Var}(\bar{y}_R|y_R) + \mathrm{Var}(\bar{y}_{NR}|y_R) - 2 \cdot \mathrm{Cov}(\bar{y}_R, \bar{y}_{NR}|y_R)$$

$$= s_R^2(1-r^{-1}) + \frac{1}{m}s_R^2(1-r^{-1}) - 2\cdot\frac{1}{m}s_R^2\cdot\text{Cov}(\bar{y}_R,\bar{y}_{NR})$$

$$\text{Var}(\bar{y}_{NR}) = (1-r^{-1})S_R^2 + \frac{1}{m}(1-\frac{r}{n})S_R^2 - 2\cdot\frac{1}{m}(1-r^{-1})S_R^2$$

$$= (1-r^{-1})S_R^2(1-m^{-1})$$

$$\text{So}\quad E\left((y_{jNR}-\bar{y}_{NR})^2\right) = \text{Var}[y_{jNR}-\bar{y}_{NR}] + E^2[y_{jNR}-\bar{y}_{NR}]$$

$$= (1-r^{-1})S_R^2(1-m^{-1})$$

$$\text{So}\quad E[S_{NR}^2] = \frac{1}{m-1}\sum_{j=1}^{m}E[(y_{jNR}-\bar{y}_{NR})^2]$$

$$= \frac{m}{m-1}(1-r^{-1})S_R^2\left(\frac{m-1}{m}\right) = (1-r^{-1})S_R^2 \quad \text{(unbiased estimator)}$$

And the third term:

$$E\left((\bar{y}_R-\bar{y}_{NR})^2|y_R\right)$$

$$= \text{Var}(\bar{y}_R-\bar{y}_{NR}|y_R)$$

$$= \text{Var}(\bar{y}_{NR}|y_R)$$

$$= \frac{1}{m}(1-r^{-1})S_R^2$$

So,

$$E[S_*^2|y_R] = E\left[\frac{(r-1)S_R^2 + (m-1)S_{NR}^2 + \frac{rm}{n}(\bar{y}_R-\bar{y}_{NR})^2}{n-1}|y_R\right]$$

$$= \frac{(r-1)}{n-1}S_R^2 + \frac{(m-1)}{n-1}E[S_{NR}^2|y_R] + \frac{rm}{n(n-1)}E[(\bar{y}_R-\bar{y}_{NR})^2|y_R]$$

$$= \frac{(r-1)}{n-1}S_R^2 + \frac{(m-1)}{n-1}(1-r^{-1})S_R^2 + \frac{rm}{n(n-1)}\cdot\frac{1}{m}(1-r^{-1})S_R^2$$

$$= S_R^2(1-r^{-1})\left[\frac{r}{n-1} + \frac{m-1}{n-1} + \frac{r}{n(n-1)}\right]$$

$$= S_R^2(1-r^{-1})\left[\frac{r+(m-1)}{n-1} + \frac{r}{n(n-1)}\right]$$

Final expression boxed:

$$\boxed{S_R^2(1-r^{-1})\left(\frac{1}{n-1}+\frac{r}{n(n-1)}+\frac{m-1}{n-1}\right)}$$

$$(\chi)=\frac{nr+n(m-1)+r}{n(n-1)}$$

$$=\frac{nr+n(n-r)-n+r}{n(n-1)}$$

$$=\frac{nr+n^2-nr-n+r}{n(n-1)}$$

$$=\frac{n^2-n+r}{n(n-1)}$$

$$=1+\frac{r}{n(n-1)}$$

Therefore:

$$E[S_*^2|y_R]=S_R^2(1-r^{-1})\cdot\left(1+\frac{r}{n(n-1)}\right)$$

(c)

Assume $N\to\infty$.

$$\bar{y}_*=\frac{r}{n}\bar{y}_R+\frac{m}{n}\bar{y}_{NR}.$$

So,

$$\text{Var}(\bar{y}_*|n,r)=\text{Var}\left(\frac{r}{n}\bar{y}_R+\frac{m}{n}\bar{y}_{NR}|n,r\right)$$

$$=\text{Var}\left(E\left[\frac{r}{n}\bar{y}_R+\frac{m}{n}\bar{y}_{NR}|y_R,n,r\right]\right)\quad\text{(Term 1)}$$

$$+E\left[\text{Var}\left(\frac{r}{n}\bar{y}_R+\frac{m}{n}\bar{y}_{NR}|y_R,n,r\right)\right].\quad\text{(Term 2)}$$

Term 1:

$$\text{Var}\left(E\left[\frac{r}{n}\bar{y}_R+\frac{m}{n}\bar{y}_{NR}|y_R,n,r\right]\right)$$

$$=\text{Var}\left(\frac{r}{n}\bar{y}_R+\frac{m}{n}E[\bar{y}_{NR}|y_R,n,r]\right)$$

$$=\text{Var}\left(\frac{r}{n}\bar{y}_R+\frac{m}{n}\bar{y}_R\right)$$

$$= \bar{y}_R^2.$$

Term 2:

$$E\left[\mathrm{Var}\left(\frac{r}{n}\bar{y}_R + \frac{m}{n}\bar{y}_{NR}|y_R, n, r\right)\right]$$

$$= E\left[\frac{m}{n^2}(1 - r^{-1})s_R^2|n, r\right]$$

$$= \frac{m}{n^2}(1 - \frac{r}{n})s_R^2|n, r.$$

$$= \frac{m}{n^2}\left(1 - \frac{r}{n}\right)E\left[\sum_{i=1}^{r}(y_{iR} - \bar{y}_R)^2|n, r\right].$$

$$E[(y_{ik} - \bar{y}_{ik})^2|n, r] = \frac{m}{n^2} \cdot \mathrm{Var}(y_{iR} - \bar{y}_{iR}|n, r)$$

$$= \frac{m}{n^2}\left(\mathrm{Var}(y_{iR}) + \mathrm{Var}(\bar{y}_{iR}) - 2 \cdot \mathrm{Cov}(y_{iR}, \bar{y}_{iR})\right)$$

$$\mathrm{Var}(y_{iR}) = r \cdot \mathrm{Var}(\bar{y}_{iR})$$

$$= \frac{m}{n^2}\left(r + 1 - 2\right) \cdot \mathrm{Var}(\bar{y}_{iR})$$

$$m = n - r \Rightarrow \frac{(n - r)(r - 1)}{n^2} \cdot \mathrm{Var}(\bar{y}_{iR})$$

$$\text{Thus} \quad \mathrm{Var}(\bar{y}_*|n, r) = \left(1 + \frac{(n - r)(r - 1)}{n^2}\right) \cdot \mathrm{Var}(\bar{y}_R).$$

$$E[(y_{ik}|n, r)] = E\left[\frac{1}{n}S_*^2|n, r\right]$$

$$= E\left[\frac{1}{n}\left((r - 1)S_R^2 + (m - 1)S_{NR}^2 + \frac{rm}{n}(\bar{y}_R - \bar{y}_{NR})^2\right)\middle| n, r\right]$$

$$= \frac{r - 1}{n(n - 1)}E[S_R^2|n, r] + \frac{m - 1}{n(n - 1)}E[S_{NR}^2|n, r] + \frac{rm}{n^2(n - 1)}E\left[(\bar{y}_R - \bar{y}_{NR})^2|n, r\right].$$

Term 1:

$$E[S_R^2|n, r]$$

$$= \frac{r}{r - 1}E\left[(y_{iR} - \bar{y}_R)^2|n, r\right]$$

$$= \frac{r}{r-1}\text{Var}[y_{iR} - \bar{y}_R | n, r]$$

$$= \frac{r}{r-1}\left(\text{Var}[y_{iR}|n,r] + \text{Var}[\bar{y}_R|n,r] - 2 \cdot \text{Cov}(y_{iR}, \bar{y}_R|n,r)\right)$$

$$= \frac{r}{r-1}\left(r \cdot \text{Var}[\bar{y}_R] + \text{Var}[\bar{y}_{NR}] - 2 \cdot \text{Var}[\bar{y}_R]\right)$$

$$= r \cdot \text{Var}[\bar{y}_R]$$

Term 2:

$$E[S^2_{NR}|n,r]$$

$$= E\left[E\left[S^2_{NR}|y_R,n,r\right]\right]$$

$$= E\left[E\left[\frac{m}{m-1}(y_{iNR} - \bar{y}_{NR})^2|y_R,n,r\right]\right]$$

$$= \frac{m}{m-1}E\left[\text{Var}(y_{iNR} - \bar{y}_{NR}|y_R,n,r)\right]$$

$$= \frac{m}{m-1}E\left[(1-\frac{r}{n})S^2_R(1-\frac{1}{m})|n,r\right]$$

$$= \frac{m}{m-1}E\left[(1-r/n)S^2_R(1-1/m)|n,r\right]$$

$$= \frac{r-1}{r} \cdot r \cdot \text{Var}(\bar{y}_R)$$

$$= (r-1)\text{Var}(\bar{y}_R)$$

Term 3:

$$E[(\bar{y}_R - \bar{y}_{NR})^2|n,r]$$

$$= E\left[E[(\bar{y}_R - \bar{y}_{NR})^2|y_R,n,r]\right]$$

$$= E\left[\frac{1}{m}(1-\frac{r}{n})S^2_R|n,r\right]$$

$$= \frac{r-1}{mr} \cdot r \cdot \text{Var}(\bar{y}_R)$$

$$E[(\bar{y}_R|n,r)] = E\left[\frac{1}{n}S_*^2|n,r\right]$$

$$= E\left[\frac{1}{n}\left((r-1)S_R^2 + (m-1)S_{NR}^2 + \frac{rm}{n}(\bar{y}_R - \bar{y}_{NR})^2\right)\middle| n,r\right]$$

$$= \frac{r-1}{n(n-1)}E[S_R^2|n,r] + \frac{m-1}{n(n-1)}E[S_{NR}^2|n,r] + \frac{rm}{n^2(n-1)}E\left[(\bar{y}_R - \bar{y}_{NR})^2|n,r\right].$$

$$= \frac{r-1}{n(n-1)} \cdot r \cdot \mathrm{Var}(\bar{y}_R) + \frac{m-1}{n(n-1)} \cdot (1 - r/n) \cdot \mathrm{Var}(\bar{y}_R) + \frac{rm}{n^2(n-1)} \cdot \mathrm{Var}(\bar{y}_R)$$

$$= \left[\frac{r-1}{n(n-1)} \cdot r + \frac{(m-1)(n-1)}{n(n-1)} + \frac{rm}{n(n-1)}\right] \cdot \mathrm{Var}(\bar{y}_R)$$

$$= \frac{r-1}{n(n-1)}(r + (m-1)(n-1) + \frac{rm}{n}) \cdot \mathrm{Var}(\bar{y}_R)$$

$$= \frac{(r-1)(n-1)(r+1)}{n(n-1)} \cdot \mathrm{Var}(\bar{y}_R)$$

$$= (r-1)\left[\frac{n-1}{n} + \frac{r}{n}\right] \cdot \mathrm{Var}(\bar{y}_R)$$

$$= \frac{(r-1)(n^2+r)}{n^2(n-1)} \cdot \mathrm{Var}(\bar{y}_R)$$

$$= \frac{(r-1)(n-1)+r}{n^2(n-1)} \cdot \mathrm{Var}(\bar{y}_R)$$

$$= \frac{(r-1)(n^2-n+r)}{n^2(n-1)} \cdot \mathrm{Var}(\bar{y}_R)$$

$$= \frac{(n-1)(n+r)}{n^2(n-1)} \cdot \mathrm{Var}(\bar{y}_R)$$

$$= \frac{(n+r)(r-1)(n+r)}{n^2(r-1)} \cdot \mathrm{Var}(\bar{y}_R)$$

$$= \frac{(n+r)(n+1)}{n^2} \cdot \mathrm{Var}(\bar{y}_R)$$

Thus,

$$\mathrm{Var}(\bar{y}_*|n,r) \geq E[(\bar{y}_R|n,r)]$$

5.11

(a)

$$E(B_*|y_R) = \frac{1}{D-1}E\left(\sum_{d=1}^{D}\left(\bar{y}_*^{(d)} - \bar{y}_*\right)^2 \Big| y_R\right)$$

$$= \frac{1}{D-1}E\left(\sum_{d=1}^{D}\bar{y}_*^{(d)2} - \frac{1}{D}\left(\sum_{d=1}^{D}\bar{y}_*^{(d)}\right)^2 \Big| y_R\right)$$

$$\forall\ d:\ E(\bar{y}_*^{(d)2}|y_R) = E(\bar{y}_*^2|y_R) = \mathrm{Var}(\bar{y}_*|y_R) + E^2(\bar{y}_*|y_R)$$

$$= (1-r^{-1})\frac{m}{n^2}S_R^2 + \bar{y}_R^2$$

$$E\left(\sum_{d=1}^{D}\bar{y}_*^{(d)2}|y_R\right) = D\cdot(1-r^{-1})\frac{m}{n^2}S_R^2 + D\cdot\bar{y}_R^2$$

$$E\left(\left(\sum_{d=1}^{D}\bar{y}_*^{(d)}\right)^2|y_R\right) = \frac{1}{n^2}E\left(\sum_{d=1}^{D}\sum_{i=1}^{n}y_i^{(d)}|y_R\right) = \frac{1}{n^2}E\left(D\cdot\sum_{i=1}^{r}y_{iR} + \sum_{j=1}^{m}y_{NR,j}\Big|y_R\right)$$

$$= \frac{Dr^2}{n^2}\bar{y}_R^2 + \frac{2Dr}{n^2}\bar{y}_R E\left(\sum_{j=1}^{m}y_{NR,j}|y_R\right) + \frac{Dm^2}{n^2}\bar{y}_R^2$$

$$= \frac{D^2r^2}{n^2}\bar{y}_R^2 + \frac{r-1}{r}\cdot\frac{Dm}{n}S_R^2$$

$$E(B_*|y_R) = \frac{1}{D-1}\left[D\cdot(1-r^{-1})\frac{m}{n^2}S_R^2 + D\bar{y}_R^2 - D\bar{y}_R^2 - \frac{r-1}{r}\cdot\frac{m}{n}S_R^2\right]$$

$$= \frac{1}{D-1}\left[D\cdot(1-r^{-1})\frac{m}{n^2}S_R^2 + D\bar{y}_R^2 - \frac{r-1}{r}\cdot\frac{m}{n^2}S_R^2\right]$$

$$= (1-r^{-1})\frac{m}{n^2}S_R^2 = \mathrm{Var}(\bar{y}_*|y_R)$$

$$\square$$

(b)

Let $\mathcal{Y}$ denote the population value.

$$\mathrm{Var}\left(\bar{y}_*|n,r,\mathcal{Y}\right) = E\left(\mathrm{Var}\left(\bar{y}_*|y_R\right)|n,r,\mathcal{Y}\right) + \mathrm{Var}\left(E\left(\bar{y}_*|y_R\right)|n,r,\mathcal{Y}\right)$$

$$E\left(\mathrm{Var}\left(\bar{y}_*|y_R\right)|n,r,\mathcal{Y}\right) = E\left(\frac{1}{D}\mathrm{Var}\left(\bar{y}_*^{(d)}|y_R\right)|n,r,\mathcal{Y}\right)$$

$$= \frac{1}{D}\mathrm{Var}\left(\bar{y}_*|n, r, \mathcal{Y}\right) - \frac{1}{D}\mathrm{Var}(\bar{y}_R|n, r, \mathcal{Y})$$

$$\mathrm{Var}\left(E\left(\bar{y}_*|y_R\right)|n, r, \mathcal{Y}\right) = \mathrm{Var}(\bar{y}_R|n, r, \mathcal{Y})$$

$$\mathrm{Var}\left(\bar{y}_*|n, r, \mathcal{Y}\right) = \frac{1}{D}\mathrm{Var}\left(\bar{y}_*|n, r, \mathcal{Y}\right) - \frac{1}{D}\mathrm{Var}(\bar{y}_R|n, r, \mathcal{Y}) + \mathrm{Var}(\bar{y}_R|n, r, \mathcal{Y})$$

$$\mathrm{Var}\left(\bar{y}_*|n, r, \mathcal{Y}\right) = \frac{1}{D}\mathrm{Var}\left(\bar{y}_*|n, r, \mathcal{Y}\right) + (1 - D^{-1})\mathrm{Var}(\bar{y}_R|n, r, \mathcal{Y})$$

$$\mathrm{Var}\left(\bar{y}_*|n, r, \mathcal{Y}\right) - \mathrm{Var}(\bar{y}_R|n, r, \mathcal{Y})$$

$$= \left(1 - D^{-1}\right)\left[\mathrm{Var}(\bar{y}_R|n, r, \mathcal{Y}) - \mathrm{Var}\left(\bar{y}_*|n, r, \mathcal{Y}\right)\right]$$

$$= \left(1 - D^{-1}\right)\left[-\frac{1}{r}(1 - \frac{r}{n})(1 - r^{-1})\mathrm{Var}\left(\bar{y}_R|n, r, \mathcal{Y}\right)\right] < 0$$

$$\mathrm{Var}(\bar{y}_*|n, r, \mathcal{Y}) - \mathrm{Var}(\bar{y}_R|n, r, \mathcal{Y}) < 0$$

So $\bar{y}_*$ is more efficient than $\bar{y}_R$