§4.4 Replication variance estimation

Replication variance estimation (under complete response)

Let $\hat{\theta}_n$ be the complete-sample estimator of θ. The replication variance estimator of $\hat{\theta}_n$ takes the form of

$$\hat{V}_{rep}(\hat{\theta}_n) = \sum_{k=1}^{L} c_k \left(\hat{\theta}_n^{(k)} - \hat{\theta}_n \right)^2 \tag{1}$$

where L is the number of replicates, c_k is the replication factor associated with replication k, and $\hat{\theta}_n^{(k)}$ is the k-th replicate of $\hat{\theta}_n$. If $\hat{\theta}_n = \sum_{i=1}^{n} y_i / n$, then we can write $\hat{\theta}_n^{(k)} = \sum_{i=1}^{n} w_i^{(k)} y_i$ for some replication weights $w_1^{(k)}, w_2^{(k)}, \cdots, w_n^{(k)}$.

Replication variance estimation (under complete response)

- For example, in the jackknife method, we have $L = n$, $c_k = (n-1)/n$, and

$$w_i^{(k)} = \begin{cases} (n-1)^{-1} & \text{if } i \neq k \\ 0 & \text{if } i = k. \end{cases}$$

If we use the above jackknife method to $\hat{\theta}_n = \sum_{i=1}^n y_i/n$, the resulting jackknife estimator in (1) is algebraically equivalent to $n^{-1}(n-1)^{-1}\sum_{i=1}^n (y_i - \bar{y}_n)^2$.

Replication variance estimation (under complete response)

- Under some regularity conditions, for $\hat{\theta}_n = g(\bar{y}_n)$, the replication variance estimator of $\hat{\theta}_n$, defined by

$$\hat{V}_{rep}\left(\hat{\theta}_n\right) = \sum_{k=1}^{L} c_k \left(\hat{\theta}_n^{(k)} - \hat{\theta}_n\right)^2, \tag{2}$$

where $\hat{\theta}_n^{(k)} = g(\bar{y}_n^{(k)})$, satisfies

$$\hat{V}_{rep}\left(\hat{\theta}_n\right) \cong \{g'(\bar{y}_n)\}^2 \hat{V}_{rep}(\bar{y}_n).$$

Thus, the replication variance estimator (2) is asymptotically equivalent to the linearized variance estimator.

Justification

Remark

- If the parameter of interest, denoted by ψ, is estimated by $\hat{\psi}$ which is obtained by solving an estimating equation $\sum_{i=1}^{n} U(\psi; y_i) = 0$, then a consistent variance estimator can be obtained by the sandwich formula: The complete-sample variance estimator of $\hat{\psi}$ is

$$\hat{V}\left(\hat{\psi}\right) = \hat{\tau}_u^{-1} \hat{\Omega}_u \hat{\tau}_u^{-1'} \tag{3}$$

where

$$\begin{aligned} \hat{\tau}_u &= n^{-1} \sum_{i=1}^{n} \dot{U}\left(\hat{\psi}; \mathbf{y}_i\right) \\ \hat{\Omega}_u &= n^{-1} (n-1)^{-1} \sum_{i=1}^{n} \left(\hat{u}_i - \bar{u}_n\right)^{\otimes 2}, \end{aligned}$$

$\dot{U}(\psi; \mathbf{y}) = \partial U(\psi; \mathbf{y}) / \partial \psi$, $\bar{u}_n = n^{-1} \sum_{i=1}^{n} \hat{u}_i$, and $\hat{u}_i = U(\hat{\psi}; \mathbf{y}_i)$.

- If we want to use the replication method of the form (1), we can construct the replication variance estimator of $\hat{\psi}$ by

$$\hat{V}_{rep}(\hat{\psi}) = \sum_{k=1}^{L} c_k \left(\hat{\psi}^{(k)} - \hat{\psi} \right)^2 \tag{4}$$

where $\hat{\psi}^{(k)}$ is computed by

$$\hat{U}^{(k)}(\psi) \equiv \sum_{i=1}^{n} w_i^{(k)} U(\psi; y_i) = 0. \tag{5}$$

- One-step approximation of $\hat{\psi}^{(k)}$ is to use

$$\hat{\psi}_1^{(k)} = \hat{\psi} - \left\{ \dot{U}^{(k)}(\hat{\psi}) \right\}^{-1} \hat{U}^{(k)}(\hat{\psi}) \tag{6}$$

or, even more simply, to use

$$\hat{\psi}_1^{(k)} = \hat{\psi} - \left\{ \dot{U}(\hat{\psi}) \right\}^{-1} \hat{U}^{(k)}(\hat{\psi}). \tag{7}$$

The replication variance estimator using (7) is algebraically equivalent to

$$\left\{ \dot{U}(\hat{\psi}) \right\}^{-1} \left[\sum_{k=1}^{n} c_k \left\{ \hat{U}^{(k)}(\hat{\psi}) - \hat{U}(\hat{\psi}) \right\}^{\otimes 2} \right] \left\{ \dot{U}(\hat{\psi}) \right\}^{-1},$$

which is very close to the sandwich variance formula in (3).

Replication variance estimation with estimated nuisance parameters

- We are interested in estimating the variance of $\hat{\psi}$ which is the solution to

$$\hat{U}_1(\psi, \hat{\eta}) \equiv E\left\{\hat{U}_{\rm com}(\psi) \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \hat{\eta}\right\} = 0$$

  where

$$\hat{U}_{\rm com}(\psi) = n^{-1}\sum_{i=1}^{n} U(\psi; \mathbf{y}_i)$$

  and $\hat{\eta}$ is the MLE of η in $p(\mathbf{y}_{\rm mis} \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \eta)$.
- The nuisance parameter η is obtained by solving

$$\hat{U}_2(\eta) \equiv E\left\{S_{\rm com}(\eta) \mid \mathbf{y}_{obs}, \boldsymbol{\delta}; \eta\right\} = 0$$

  where

$$S_{\rm com}(\eta) = n^{-1}\sum_{i=1}^{n} S(\eta; \mathbf{y}_i, \boldsymbol{\delta}_i)$$

- Thus, $(\hat{\psi}, \hat{\eta})$ is the solution to the joint estimating equation:

$$\begin{pmatrix} \hat{U}_1(\psi, \eta) \\ \hat{U}_2(\eta) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- For variance estimation of $\hat{\psi}$, we can either use linearization or replication.
- For the replication method, we obtain $(\hat{\psi}^{(k)}, \hat{\eta}^{(k)})$ by solving the replicated estimating equation

$$\begin{pmatrix} \hat{U}_1^{(k)}(\psi, \eta) \\ \hat{U}_2^{(k)}(\eta) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

where

$$\hat{U}_1^{(k)}(\psi, \eta) \equiv \sum_{i=1}^{n} w_i^{(k)} E\left\{ U(\psi; \mathbf{y}_i) \mid \mathbf{y}_{\mathrm{obs,i}}, \boldsymbol{\delta}_i; \eta \right\}$$

and

$$\hat{U}_2^{(k)}(\eta) \equiv \sum_{i=1}^{n} w_i^{(k)} E\left\{ S(\eta; \mathbf{y}_i, \boldsymbol{\delta}_i) \mid \mathbf{y}_{obs,i}, \boldsymbol{\delta}_i; \eta \right\}$$

Back to Example 4.1

- For the regression imputation in Example 4.1,

$$\hat{\theta}_{Id} = \frac{1}{n}\sum_{i=1}^{n}\left\{\delta_i y_i + (1-\delta_i)\left(\hat{\beta}_0 + \hat{\beta}_1 x_i\right)\right\}.$$

- The replication variance estimator of $\hat{\theta}_{Id}$ is computed by

$$\hat{V}_{rep}\left(\hat{\theta}_{Id}\right) = \sum_{k=1}^{L} c_k \left(\hat{\theta}_{Id}^{(k)} - \hat{\theta}_{Id}\right)^2 \tag{8}$$

where

$$\hat{\theta}_{Id}^{(k)} = \sum_{i=1}^{n} w_i^{(k)}\left\{\delta_i y_i + (1-\delta_i)\left(\hat{\beta}_0^{(k)} + \hat{\beta}_1^{(k)} x_i\right)\right\}$$

and $(\hat{\beta}_0^{(k)}, \hat{\beta}_1^{(k)})$ is the solution to

$$\sum_{i=1}^{n} w_i^{(k)}\delta_i\left(y_i - \beta_0 - \beta_1 x_i\right)(1, x_i) = (0,0).$$

Example 4.4

- We now return to the setup of Example 3.11.
- In this case, the deterministically imputed estimator of $\theta = E(Y)$ is constructed by

$$\hat{\theta}_{Id} = n^{-1} \sum_{i=1}^{n} \{\delta_i y_i + (1-\delta_i)\hat{p}_{0i}\} \tag{9}$$

where $\hat{p}_{0i}$ is the predictor of y_i given x_i and $\delta_i = 0$. That is,

$$\hat{p}_{0i} = \frac{p(x_i;\hat{\beta})\{1-\pi(x_i,1;\hat{\phi})\}}{\{1-p(x_i;\hat{\beta})\}\{1-\pi(x_i,0;\hat{\phi})\} + p(x_i;\hat{\beta})\{1-\pi(x_i,1;\hat{\phi})\}},$$

where $\hat{\beta}$ and $\hat{\phi}$ are jointly estimated by the EM algorithm.

Example 4.4 (Cont'd)

E-step

$$\bar{S}_1\left(\beta \mid \beta^{(t)}, \phi^{(t)}\right) = \sum_{\delta_i=1}\{y_i - p_i(\beta)\}\mathbf{x}_i + \sum_{\delta_i=0}\sum_{j=0}^{1} w_{ij(t)}\{j - p_i(\beta)\}\mathbf{x}_i,$$

where

$$\begin{aligned} w_{ij(t)} &= Pr\left(Y_i = j \mid \mathbf{x}_i, \delta_i = 0; \beta^{(t)}, \phi^{(t)}\right) \\ &= \frac{Pr\left(Y_i = j \mid \mathbf{x}_i; \beta^{(t)}\right) Pr\left(\delta_i = 0 \mid \mathbf{x}_i, j; \phi^{(t)}\right)}{\sum_{y=0}^{1} Pr\left(Y_i = y \mid \mathbf{x}_i; \beta^{(t)}\right) Pr\left(\delta_i = 0 \mid \mathbf{x}_i, y; \phi^{(t)}\right)} \end{aligned}$$

and

$$\begin{aligned} \bar{S}_2\left(\phi \mid \beta^{(t)}, \phi^{(t)}\right) &= \sum_{\delta_i=1}\{\delta_i - \pi(\mathbf{x}_i, y_i; \phi)\}(\mathbf{x}_i', y_i)' \\ &\quad + \sum_{\delta_i=0}\sum_{j=0}^{1} w_{ij(t)}\{\delta_i - \pi_i(\mathbf{x}_i, j; \phi)\}(\mathbf{x}_i', j)'. \end{aligned}$$

Example 4.4 (Cont'd)

M-step

The parameter estimates are updated by solving

$$\left[\bar{S}_1\left(\beta \mid \beta^{(t)}, \phi^{(t)}\right), \bar{S}_2\left(\phi \mid \beta^{(t)}, \phi^{(t)}\right)\right] = (0, 0)$$

for β and ϕ.

Example 4.4 (Cont'd)

- For replication variance estimation, we can use (8) with

$$\hat{\theta}_{Id}^{(k)} = \sum_{i=1}^{n} w_i^{(k)} \left\{ \delta_i y_i + (1 - \delta_i) \hat{p}_{0i}^{(k)} \right\}. \tag{10}$$

where

$$\hat{p}_{0i}^{(k)} = \frac{p(x_i; \hat{\beta}^{(k)})\{1 - \pi(x_i, 1; \hat{\phi}^{(k)})\}}{\{1 - p(x_i; \hat{\beta}^{(k)})\}\pi(x_i, 0; \hat{\phi}^{(k)}) + p(x_i; \hat{\beta}^{(k)})\{1 - \pi(x_i, 1; \hat{\phi}^{(k)})\}}.$$

and $(\hat{\beta}^{(k)}, \hat{\phi}^{(k)})$ is obtained by solving the mean score equations with weights replaced by the replication weights $w_i^{(k)}$.

Example 4.4 (Cont'd)

- That is, $(\hat{\beta}^{(k)}, \hat{\phi}^{(k)})$ is the solution to

$$
\begin{aligned}
\bar{S}_1^{(k)}(\beta, \phi) &\equiv \sum_{\delta_i=1} w_i^{(k)} \{y_i - p(\mathbf{x}_i; \beta)\} \mathbf{x}_i \\
&\quad + \sum_{\delta_i=0} w_i^{(k)} \sum_{y=0}^{1} w_{iy}^*(\beta, \phi)\{y - p(\mathbf{x}_i; \beta)\}\mathbf{x}_i = 0 \\
\bar{S}_2^{(k)}(\beta, \phi) &\equiv \sum_{\delta_i=1} w_i^{(k)} \{\delta_i - \pi(\mathbf{x}_i, y_i; \phi)\} (\mathbf{x}_i', y_i)' \\
&\quad + \sum_{\delta_i=0} w_i^{(k)} \sum_{y=0}^{1} w_{iy}^*(\beta, \phi)\{\delta_i - \pi(\mathbf{x}_i, y; \beta)\}(\mathbf{x}_i', y)' = 0
\end{aligned}
$$

and

$$
w_{iy}^*(\beta, \phi) = \frac{p(\mathbf{x}_i; \beta)\pi(\mathbf{x}_i, 1; \phi)}{\{1 - p(\mathbf{x}_i; \beta)\}\pi(\mathbf{x}_i, 0; \phi) + p(\mathbf{x}_i; \beta)\pi(\mathbf{x}_i, 1; \phi)}.
$$