4.2 Basic Theory for Imputation

Basic Setup

- Recall that imputation is used in computing the expected estimating equation

$$n^{-1}\sum_{i=1}^{n} U(\psi;\mathbf{y}_i)=0.$$

- $\mathbf{y}_{\rm mis}^{*(1)},\cdots,\mathbf{y}_{\rm mis}^{*(m)}$: m imputed values of $\mathbf{y}_{\rm mis}$ generated from

$$f(\mathbf{y}_{\rm mis}\mid \mathbf{y}_{\rm obs},\boldsymbol{\delta};\hat{\eta}_p)=\frac{f(\mathbf{y};\hat{\theta}_p)P(\boldsymbol{\delta}\mid \mathbf{y};\hat{\phi}_p)}{\int f(\mathbf{y};\hat{\theta}_p)P(\boldsymbol{\delta}\mid \mathbf{y};\hat{\phi}_p)d\mu(\mathbf{y}_{\rm mis})}, \qquad (1)$$

  where $\hat{\eta}_p=(\hat{\theta}_p,\hat{\phi}_p)$ is a preliminary estimator of $\eta_0=(\theta_0,\phi_0)$.
- Two cases for ψ:
  1. $\psi=\eta$: Theorem 4.1 (i.e. U is the score function for η)
  2. $\psi\neq\eta$: Theorem 4.2 (for general estimating function U)

Case 1: $\psi = \eta$

- Using m imputed values, the imputed score function is computed as

$$\bar{S}_{I,m}(\eta \mid \hat{\eta}_p) \equiv \frac{1}{m}\sum_{j=1}^{m} S_{\rm com}(\eta; \mathbf{y}_{obs}, \mathbf{y}_{mis}^{*(j)}, \boldsymbol{\delta}),$$

  where $S_{\rm com}(\eta; \mathbf{y}) = n^{-1}\sum_{i=1}^{n} S(\eta; \mathbf{y}_i)$ is the score function of $\eta = (\theta, \phi)$ under complete response.

- We will first consider the case when $m \to \infty$. In this case,

$$\lim_{m\to\infty} \bar{S}_{I,m}(\eta \mid \hat{\eta}_p) = E\left\{S_{\rm com}(\eta) \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \hat{\eta}_p\right\} := \bar{S}(\eta \mid \hat{\eta}_p).$$

  The solution to $\bar{S}(\eta \mid \hat{\eta}_p) = 0$ is the M-step update of the EM algorithm using $\hat{\eta}_p$ as the current parameter estimate.

Preliminary results

- Let $\hat{\eta}_1$ and $\hat{\eta}_2$ satisfy
$$\bar{S}(\hat{\eta}_2 \mid \hat{\eta}_1) = 0$$
where
$$\bar{S}\left(\eta_2 \mid \eta_1\right) = \int S_{\rm com}(\eta_2) f(\mathbf{y}_{\rm mis} \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \eta_1) d\mathbf{y}_{\rm mis}.$$
- To apply Taylor expansion, we need to use
$$\frac{\partial}{\partial \eta_2'} \bar{S}(\eta_2 \mid \eta_1) = \int \left\{ \frac{\partial}{\partial \eta_2'} S_{\rm com}(\eta_2) \right\} f(\mathbf{y}_{\rm mis} \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \eta_1) d\mathbf{y}_{\rm mis}$$
and
$$\frac{\partial}{\partial \eta_1'} \bar{S}(\eta_2 \mid \eta_1) = \int S_{\rm com}(\eta_2) S_{\rm mis}(\eta_1)' f(\mathbf{y}_{\rm mis} \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \eta_1) d\mathbf{y}_{\rm mis}.$$

Theory for Case 1 with $m = \infty$

Lemma 4.2 (Asymptotic results for $m = \infty$)

Assume that $\hat{\eta}_p$ converges in probability to η_0. Let $\hat{\eta}_{I,m}$ be the solution to

$$\frac{1}{m}\sum_{j=1}^{m} S_{\text{com}}\left(\eta; \mathbf{y}_{obs}, \mathbf{y}_{mis}^{*(j)}, \boldsymbol{\delta}\right) = 0,$$

where $\mathbf{y}_{mis}^{*(1)}, \cdots, \mathbf{y}_{mis}^{*(m)}$ are the imputed values generated from $f(\mathbf{y}_{\text{mis}} \mid \mathbf{y}_{\text{obs}}, \boldsymbol{\delta}; \hat{\eta}_p)$. Then, under some regularity conditions, for $m \to \infty$,

$$\hat{\eta}_{I,\infty} \cong \hat{\eta}_{\text{MLE}} + \mathcal{J}_{\text{mis}}\left(\hat{\eta}_p - \hat{\eta}_{\text{MLE}}\right) \tag{2}$$

and

$$V\left(\hat{\eta}_{I,\infty}\right) \doteq \mathcal{I}_{obs}^{-1} + \mathcal{J}_{mis}\left\{V\left(\hat{\eta}_p\right) - V\left(\hat{\eta}_{\text{MLE}}\right)\right\}\mathcal{J}_{mis}', \tag{3}$$

where $\mathcal{J}_{mis} = \mathcal{I}_{com}^{-1}\mathcal{I}_{mis}$ is the fraction of missing information.

Remark 1

- Equation (2) means that

$$\hat{\eta}_{I,\infty} = \left(I - \mathcal{J}_{\text{mis}}\right) \hat{\eta}_{MLE} + \mathcal{J}_{\text{mis}} \hat{\eta}_p. \tag{4}$$

  That is, $\hat{\eta}_{I,\infty}$ is a convex combination of $\hat{\eta}_{MLE}$ and $\hat{\eta}_p$.
- Note that $\hat{\eta}_{I,\infty}$ is one-step EM update with initial estimate $\hat{\eta}_p$. Let $\hat{\eta}^{(t)}$ be the t-th EM update of η that is computed by solving

$$\bar{S}\left(\eta \mid \hat{\eta}^{(t-1)}\right) = 0$$

  with $\hat{\eta}^{(0)} = \hat{\eta}_p$. Equation (4) implies that

$$\hat{\eta}^{(t)} = \left(I - \mathcal{J}_{\text{mis}}\right) \hat{\eta}_{MLE} + \mathcal{J}_{\text{mis}} \hat{\eta}^{(t-1)}. \tag{5}$$

- By (5), we can obtain

$$\hat{\eta}^{(t)} = \hat{\eta}_{MLE} + \left(\mathcal{J}_{\text{mis}}\right)^{t-1} \left(\hat{\eta}^{(0)} - \hat{\eta}_{MLE}\right),$$

which justifies $\lim_{t\to\infty} \hat{\eta}^{(t)} = \hat{\eta}_{MLE}$.

- Also, (5) implies that the convergence rate for EM algorithm is linear. It can be shown that

$$\eta^{(t+1)} - \eta^{(t)} \cong \mathcal{J}_{\text{mis}} \left(\eta^{(t)} - \eta^{(t-1)}\right) \tag{6}$$

where $\mathcal{J}_{\text{mis}} = \mathcal{I}_{\text{com}}^{-1}\mathcal{I}_{\text{mis}}$ is called the *fraction of missing information*. The fraction of missing information may vary across different components of $\eta^{(t)}$, suggesting that certain components of $\eta^{(t)}$ may approach η^{*} rapidly while other components may require many iterations. Roughly speaking, the rate of convergence of a vector sequence $\eta^{(t)}$ from the EM algorithm is given by the largest eigenvalue of the matrix $\mathcal{J}_{\text{mis}}$.

Remark 2

- To obtain (3), we use the following (approximate) equality

$$V(\hat{\eta}_p) = V(\hat{\eta}_{MLE}) + V(\hat{\eta}_p - \hat{\eta}_{MLE}) \tag{7}$$

  which is essentially a version of Pythagorian theorem.
- To prove (7), we can use the following result: Any consistent and optimal estimator of θ is uncorrelated with any estimator with zero mean.

Proof for Lemma 4.2

Alternative Proof

- The solution $\hat{\eta}$ to $S_{\rm com}(\eta) = 0$ satisfies

$$\hat{\eta} - \eta_0 \cong \mathcal{I}_{\rm com}^{-1} S_{\rm com}(\eta_0).$$

- Similarly, the solution $\hat{\eta}_{I,\infty}$ to $E\{S_{\rm com}(\eta) \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \hat{\eta}_p\} = 0$ satisfies

$$\hat{\eta}_{I,\infty} - \eta_0 \cong \mathcal{I}_{\rm com}^{-1} E\{S_{\rm com}(\eta_0) \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \hat{\eta}_p\}. \tag{8}$$

- Now, we can decompose

$$E\{S_{\rm com}(\eta_0) \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \hat{\eta}_p\} = S_{\rm obs}(\eta_0) + E\{S_{\rm mis}(\eta_0) \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \hat{\eta}_p\}$$

and use

$$S_{\rm obs}(\eta_0) \cong \mathcal{I}_{\rm obs}\left(\hat{\eta}_{\rm MLE} - \eta_0\right)$$

- If $\hat{\eta}_p$ converges in probability to η_0, we can obtain

$$E\{S_{\rm mis}(\eta_0) \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \hat{\eta}_p\} \cong \mathcal{I}_{\rm mis}\left(\hat{\eta}_p - \eta_0\right).$$

- Combining the two, we obtain

$$\hat{\eta}_{I,\infty} - \eta_0 \cong \mathcal{I}_{\rm com}^{-1}\mathcal{I}_{\rm obs}\left(\hat{\eta}_{\rm MLE} - \eta_0\right) + \mathcal{I}_{\rm com}^{-1}\mathcal{I}_{\rm mis}\left(\hat{\eta}_p - \eta_0\right)$$

which essentially proves (2).

Lemma 4.3: Asymptotic results for $m = 1$

Let $S_I^*(\eta \mid \hat{\eta}_p) = S_{\rm com}(\eta; \mathbf{y}^*)$ be the imputed score function evaluated with $\mathbf{y}^* = (\mathbf{y}_{\rm obs}, \mathbf{y}^*_{\rm mis})$ where $\mathbf{y}^*_{\rm mis}$ is generated from $f\left(\mathbf{y}_{\rm mis} \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \hat{\eta}_p\right)$. Assume that $\hat{\eta}_p$ converges in probability to η_0. Then, under some regularity conditions, the solution $\hat{\eta}^*$ to $S_I^*(\eta \mid \hat{\eta}_p) = 0$ satisfies

$$\hat{\eta}^* \cong \hat{\eta}_{MLE} + \mathcal{J}_{\rm mis}\left(\hat{\eta}_p - \hat{\eta}_{MLE}\right) + \mathcal{I}_{\rm com}^{-1} S^*_{\rm mis}(\hat{\eta}_p) \tag{9}$$

where $\mathcal{J}_{\rm mis} = \mathcal{I}_{\rm com}^{-1}\mathcal{I}_{\rm mis}$ and $S^*_{\rm mis}(\eta) = S_{\rm com}(\eta; \mathbf{y}^*) - S_{\rm obs}(\eta)$.

Comparing (9) with (2), we have an additional term $\mathcal{I}_{\rm com}^{-1} S^*_{\rm mis}(\hat{\eta}_p)$ in (9) due to the imputation mechanism (i.e., the randomness in generating $\mathbf{y}^*_{\rm mis}$ from $f(\mathbf{y}_{\rm mis} \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \hat{\eta}_p)$).

Proof

- Let us write

$$S_{\mathrm{com}}(\eta; \mathbf{y}_{\mathrm{obs}}, \mathbf{y}^*_{\mathrm{mis}}) = E^*\{S_{\mathrm{com}}(\eta) \mid \mathbf{y}_{\mathrm{obs}}, \boldsymbol{\delta}; \hat{\eta}_p\}$$

and let $\hat{\eta}_I^*$ be the solution to $E^*\{S_{\mathrm{com}}(\eta) \mid \mathbf{y}_{\mathrm{obs}}, \boldsymbol{\delta}; \hat{\eta}_p\} = 0$.

- Similarly to (8), we obtain

$$\hat{\eta}_I^* - \eta_0 \cong \mathcal{I}_{\mathrm{com}}^{-1} E^*\{S_{\mathrm{com}}(\eta_0) \mid \mathbf{y}_{\mathrm{obs}}, \boldsymbol{\delta}; \hat{\eta}_p\}.$$

and we can decompose

$$E^*\{S_{\mathrm{com}}(\eta_0) \mid \mathbf{y}_{\mathrm{obs}}, \boldsymbol{\delta}; \hat{\eta}_p\} = S_{\mathrm{obs}}(\eta_0) + E\{S_{\mathrm{mis}}(\eta_0) \mid \mathbf{y}_{\mathrm{obs}}, \boldsymbol{\delta}; \hat{\eta}_p\} + Z^*$$

where

$$Z^* = E^*\{S_{\mathrm{com}}(\eta_0) \mid \mathbf{y}_{\mathrm{obs}}, \boldsymbol{\delta}; \hat{\eta}_p\} - E\{S_{\mathrm{com}}(\eta_0) \mid \mathbf{y}_{\mathrm{obs}}, \boldsymbol{\delta}; \hat{\eta}_p\}.$$

- Thus,
$$\hat{\eta}^* \cong \mathcal{I}_{\rm com}^{-1}\mathcal{I}_{\rm obs}\hat{\eta}_{\rm MLE} + \mathcal{I}_{\rm com}^{-1}\mathcal{I}_{\rm mis}\hat{\eta}_p + \mathcal{I}_{\rm com}^{-1}Z^*.$$
- Also, conditional on $(\mathbf{y}_{\rm obs}, \boldsymbol{\delta})$ and $\hat{\eta}_p$, Z^* is a random variable with mean zero and its variance
$$V\left\{S_{\rm com}(\eta_0; \mathbf{y}_{\rm obs}, \mathbf{y}_{\rm mis}) \mid \mathbf{y}_{obs}, \hat{\eta}_p\right\} = I_{\rm mis}(\hat{\eta}_p)$$
which will converge to $\mathcal{I}_{\rm mis}$ in probability.
- Therefore, we can express
$$\hat{\eta}^* \cong \mathcal{I}_{\rm com}^{-1}\mathcal{I}_{\rm obs}\left(\hat{\eta}_{\rm MLE} - \eta_0\right) + \mathcal{I}_{\rm com}^{-1}\mathcal{I}_{\rm mis}\left(\hat{\eta}_p - \eta_0\right) + \mathcal{I}_{\rm com}^{-1}Z^* \qquad (10)$$
where $Z^* \mid (\mathbf{y}_{obs}, \hat{\eta}_p) \sim (0, I_{\rm mis}(\hat{\eta}_p))$.

Main Theory: Wang and Robins (1998)

Theorem 4.1

Let $\hat{\eta}_p$ be a preliminary $\sqrt{n}$-consistent estimator of η_0 with variance V_p. Under some regularity conditions, the solution $\hat{\eta}_{I,m}$ to

$$\bar{S}_{I,m}(\eta \mid \hat{\eta}_p) \equiv \frac{1}{m}\sum_{j=1}^{m} S_{\mathrm{com}}\left(\eta; \mathbf{y}_{obs}, \mathbf{y}_{mis}^{*(j)}, \boldsymbol{\delta}\right) = 0$$

Now, the solution $\hat{\eta}_{I,m}$ to $\bar{S}_{I,m}(\eta \mid \hat{\eta}_p) = 0$ satisfies

$$\hat{\eta}_{I,m} \cong \hat{\eta}_{MLE} + \mathcal{J}_{\mathrm{mis}}(\hat{\eta}_p - \hat{\eta}_{MLE}) + \mathcal{I}_{\mathrm{com}}^{-1} \cdot \frac{1}{m}\sum_{k=1}^{m} Z^{*(k)}, \tag{11}$$

where

$$Z^{*(k)} \mid (\mathbf{y}_{obs}, \hat{\eta}_p) \overset{i.i.d}{\sim} (0, \mathcal{I}_{mis}).$$

Remark

- Theorem 4.1 implies that $\hat{\eta}_{I,m}$ is asymptotically unbiased to η and the asymptotic variance equal to

$$V\left(\hat{\eta}_{I,m}\right) \doteq \mathcal{I}_{\text{obs}}^{-1} + \mathcal{J}_{\text{mis}}\left\{V_p - \mathcal{I}_{\text{obs}}^{-1}\right\}\mathcal{J}_{\text{mis}}' + m^{-1}\mathcal{I}_{\text{com}}^{-1}\mathcal{I}_{\text{mis}}\mathcal{I}_{\text{com}}^{-1} \tag{12}$$

  where $\mathcal{J}_{\text{mis}} = \mathcal{I}_{\text{com}}^{-1}\mathcal{I}_{\text{mis}}$.
- Variance term (12) has three components. Writing

$$\hat{\eta}_{I,m} = \hat{\eta}_{MLE} + \left(\hat{\eta}_{I,\infty} - \hat{\eta}_{MLE}\right) + \left(\hat{\eta}_{I,m} - \hat{\eta}_{I,\infty}\right),$$

  we can establish that the three terms are independent and satisfies

$$\begin{aligned} V\left(\hat{\eta}_{MLE}\right) &= \mathcal{I}_{\text{obs}}^{-1}, \\ V\left(\hat{\eta}_{I,\infty} - \hat{\eta}_{MLE}\right) &= \mathcal{J}_{\text{mis}}\left\{V_p - \mathcal{I}_{\text{obs}}^{-1}\right\}\mathcal{J}_{\text{mis}}', \\ V\left(\hat{\eta}_{I,m} - \hat{\eta}_{I,\infty}\right) &= m^{-1}\mathcal{I}_{\text{com}}^{-1}\mathcal{I}_{\text{mis}}\mathcal{I}_{\text{com}}^{-1}. \end{aligned}$$

Basic Setup for Case 2 ($\psi \neq \eta$)

- Parameter ψ defined by $E\{U(\psi; \mathbf{y})\} = 0$.
- Under complete response, a consistent estimator of ψ can be obtained by solving $\hat{U}(\psi; \mathbf{y}) = 0$, where $\hat{U}(\psi; \mathbf{y}) = n^{-1} \sum_{i=1}^{n} U(\psi; \mathbf{y}_i)$.
- Assume that some part of $\mathbf{y}$, denoted by $\mathbf{y}_{\rm mis}$, is not observed and m imputed values, say $\mathbf{y}_{\rm mis}^{*(1)}, \cdots, \mathbf{y}_{\rm mis}^{*(m)}$, are generated from $f(\mathbf{y}_{\rm mis} \mid \mathbf{y}_{\rm obs}, \boldsymbol{\delta}; \hat{\eta}_{MLE})$, where $\hat{\eta}_{MLE}$ is the MLE of η_0.
- The imputed estimating function using m imputed values is computed as

$$\bar{U}_m(\psi \mid \hat{\eta}_{MLE}) = \frac{1}{m} \sum_{j=1}^{m} \hat{U}(\psi; \mathbf{y}^{*(j)}), \tag{13}$$

where $\mathbf{y}^{*(j)} = (\mathbf{y}_{\rm obs}, \mathbf{y}_{\rm mis}^{*(j)})$.

Robins and Wang (2000)

Theorem 4.2

Suppose that the parameter of interest ψ_0 is estimated by solving $\hat{U}(\psi) = 0$ under complete response. Then, under some regularity conditions, the solution to

$$E\left\{\hat{U}(\psi) \mid \mathbf{y}_{\text{obs}}, \boldsymbol{\delta}; \hat{\eta}_{MLE}\right\} = 0 \tag{14}$$

has mean ψ_0 and the asymptotic variance $\tau^{-1}\Omega\tau'^{-1}$, where

$$\begin{aligned} \tau &= -E\left\{\partial\hat{U}(\psi_0)/\partial\psi'\right\} \\ \Omega &= V\left\{\bar{U}(\psi_0 \mid \eta_0) + \kappa^* S_{\text{obs}}(\eta_0)\right\} \end{aligned}$$

and

$$\kappa^* = E\left\{\hat{U}(\psi_0) S_{\text{mis}}(\eta_0)'\right\}\mathcal{I}_{\text{obs}}^{-1}. \tag{15}$$

Remark

- To understand (15), we first consider

$$\bar{U}_{\kappa}(\psi \mid \mathbf{y}_{\text{obs}}, \boldsymbol{\delta}; \eta) = E\left\{\hat{U}(\psi) \mid \mathbf{y}_{\text{obs}}, \boldsymbol{\delta}; \eta\right\} + \kappa S_{\text{obs}}(\eta).$$

- Note that, since $S_{obs}(\hat{\eta}_{MLE}) = 0$, we have

$$\bar{U}_{\kappa}(\psi \mid \mathbf{y}_{\text{obs}}, \boldsymbol{\delta}; \hat{\eta}_{MLE}) = E\left\{\hat{U}(\psi) \mid \mathbf{y}_{\text{obs}}, \boldsymbol{\delta}; \hat{\eta}_{MLE}\right\}$$

  for all κ.
- The particular choice of $\kappa = \kappa^*$ in (15) is obtained from

$$E\left\{\frac{\partial}{\partial \eta'}\bar{U}_{\kappa}(\psi_0 \mid \mathbf{y}_{\text{obs}}, \boldsymbol{\delta}; \eta_0)\right\} = 0. \tag{16}$$

  Thus, the estimation error of $\hat{\eta}_{MLE}$ in $\bar{U}_{\kappa}(\psi \mid \mathbf{y}_{\text{obs}}, \boldsymbol{\delta}; \hat{\eta}_{MLE})$ can be safely ignored at $\kappa = \kappa^*$. Equation (16) is often called Randles (1982) condition.

Sketched Proof

- Writing
$$\bar{U}(\psi \mid \eta) \equiv E\{\hat{U}(\psi) \mid \mathbf{y}_{obs}, \boldsymbol{\delta}; \eta\},$$
the solution to (14) can be treated as the solution to the joint estimating equation
$$\mathbf{U}(\psi, \eta) \equiv \begin{bmatrix} \hat{U}_1(\psi, \eta) \\ \hat{U}_2(\eta) \end{bmatrix} = \mathbf{0},$$
where $\hat{U}_1(\psi, \eta) = \bar{U}(\psi \mid \eta)$ and $\hat{U}_2(\eta) = S_{\text{obs}}(\eta)$.
- We can apply the Taylor expansion to get
$$\begin{pmatrix} \hat{\psi} \\ \hat{\eta} \end{pmatrix} \cong \begin{pmatrix} \psi_0 \\ \eta_0 \end{pmatrix} - \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}^{-1} \begin{bmatrix} \hat{U}_1(\psi_0, \eta_0) \\ \hat{U}_2(\eta_0) \end{bmatrix}$$
where
$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{bmatrix} E\left(\partial \hat{U}_1 / \partial \psi'\right) & E\left(\partial \hat{U}_1 / \partial \eta'\right) \\ E\left(\partial \hat{U}_2 / \partial \psi'\right) & E\left(\partial \hat{U}_2 / \partial \eta'\right) \end{bmatrix}.$$

Sketched Proof (Cont'd)

- Note that

$$
\begin{aligned}
B_{11} &= E\{\partial \hat{U}(\psi_0)/\partial \psi'\} \\
B_{21} &= 0 \\
B_{12} &= E\{\hat{U}(\psi_0) S_{mis}(\eta_0)'\} \\
B_{22} &= -\mathcal{I}_{\text{obs}}
\end{aligned}
$$

- Thus,

$$
\hat{\psi} \cong \psi_0 - B_{11}^{-1} \left\{ \hat{U}_1(\psi_0, \eta_0) - B_{12} B_{22}^{-1} \hat{U}_2(\eta_0) \right\}.
$$

REFERENCES

Randles, R. H. (1982), 'On the asymptotic normality of statistics with estimated parameters', *Annals of Statistics* **10**, 462–474.

Robins, J. M. and N. Wang (2000), 'Inference for imputation estimators', *Biometrika* **87**, 113–124.

Wang, N. and J. M. Robins (1998), 'Large-sample theory for parametric multiple imputation procedures', *Biometrika* **85**, 935–948.