Optimal Designs for the Prediction of Individual Parameters in Multiple Group Random Coefficient Regression Models

Maryna Prus

Deutsche Forschungsgemeinschaft

September 8, 2017

Otto von Guericke Universität Magdeburg

Contents

Two groups random coefficient model

Group 1 (G1) - treatment - n_1 individuals
Group 2 (G2) - control - n_2 individuals

$$\text{G1}: \quad Y_{1ik} = \alpha_i + \mu_i + \varepsilon_{1ik}, \quad i = i_1, .., i_{n_1}, \quad k = 1, .., K$$

$$\text{G2}: \quad Y_{2ik} = \mu_i + \varepsilon_{2ik}, \quad i = i_{n_1+1}, .., i_N, \quad k = 1, .., K$$

$$N = n_1 + n_2$$

- K number of observations per individual
- ε_{jik}, $j = 1, 2$, observational errors
  - $\mathrm{E}(\varepsilon_{jik}) = 0$
  - $\mathrm{Var}(\varepsilon_{jik}) = \sigma^2$

Two groups random coefficient model

$$\text{G1}: \quad Y_{1ik} = \alpha_i + \mu_i + \varepsilon_{1ik}, \quad i = i_1, .., i_{n_1}, \quad k = 1, .., K$$

$$\text{G2}: \quad Y_{2ik} = \mu_i + \varepsilon_{2ik}, \quad i = i_{n_1+1}, .., i_N, \quad k = 1, .., K$$

- $\boldsymbol{\theta}_i := (\alpha_i, \mu_i)^\top$ *individual random parameters*
  - $\mathrm{E}(\boldsymbol{\theta}_i) = (\alpha_0, \mu_0)^\top =: \boldsymbol{\theta}_0$ *unknown*
  - $\mathrm{Cov}(\boldsymbol{\theta}_i) = \sigma^2 \begin{pmatrix} v & 0 \\ 0 & u \end{pmatrix}$; v & u *known*; $v,\ u > 0$
- *All* $\boldsymbol{\theta}_i$ *and all* $\varepsilon_{ji'k}$ *uncorrelated*

Search for optimal group sizes for *a) estimation of* α_0
b) prediction of $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)^\top$

Best estimation

Best linear unbiased estimator for population parameters $\boldsymbol{\theta}_0$:

$$\hat{\boldsymbol{\theta}}_0 = (\hat{\alpha}_0, \hat{\mu}_0)^\top$$

$$\hat{\alpha}_0 = \bar{Y}_1 - \bar{Y}_2 \qquad \& \qquad \hat{\mu}_0 = \bar{Y}_2$$

$$\bar{Y}_j = \frac{1}{n_j}\sum_{i=1}^{n_j}\frac{1}{K}\sum_{i=1}^{K} Y_{jik}, \quad j = 1, 2$$

Variance of estimator $\hat{\alpha}_0$:

$$\mathrm{var}\,(\hat{\alpha}_0) = \frac{\sigma^2}{K}\left(\frac{K(u+v)+1}{n_1} + \frac{Ku+1}{n_2}\right)$$

Best prediction

Best linear unbiased predictor for individual parameters $\boldsymbol{\theta}_i$:

$$\hat{\boldsymbol{\theta}}_i = (\hat{\alpha}_i, \hat{\mu}_i)^\top$$

$$\hat{\alpha}_i = \begin{cases} \frac{Kv}{K(v+u)+1}\,(\bar{Y}_{1i} - \bar{Y}_2) + \frac{Ku+1}{K(v+u)+1}\,(\bar{Y}_1 - \bar{Y}_2), & \text{ind. "}i\text{" in G1} \\ \bar{Y}_1 - \bar{Y}_2, & \text{ind. "}i\text{" in G2} \end{cases}$$

$$\bar{Y}_{ji} = \frac{1}{K}\sum_{k=1}^{K} Y_{jik}, \quad j = 1, 2$$

$$\hat{\mu}_i = \begin{cases} \frac{Ku}{K(v+u)+1}\,(\bar{Y}_{1i} - \bar{Y}_1) + \bar{Y}_2, & \text{ind. "}i\text{" in G1} \\ \frac{Ku}{Ku+1}\,\bar{Y}_{2i} + \frac{1}{Ku+1}\,\bar{Y}_2, & \text{ind. "}i\text{" in G2} \end{cases}$$

MSE matrix

MSE matrix of predictor $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_1, ..., \hat{\alpha}_N)^\top$:

$$\mathrm{Cov}\left(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\right) = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

$$\mathbf{A}_{11} = \sigma^2(Ku+1)\left(\left(\frac{N}{Kn_1 n_2} - \frac{v}{K(u+v)+1}\right)\mathbf{1}_{n_1}\mathbf{1}_{n_1}^\top + \frac{v}{K(u+v)+1}\,\mathbf{I}_{n_1}\right)$$

$$\mathbf{A}_{12} = \sigma^2(Ku+1)\frac{N}{Kn_1 n_2}\mathbf{1}_{n_1}\mathbf{1}_{n_2}^\top = \mathbf{A}_{21}^\top$$

$$\mathbf{A}_{22} = \sigma^2\left(\left(\frac{K(u+v)+1}{Kn_1} + \frac{Ku+1}{Kn_2}\right)\mathbf{1}_{n_2}\mathbf{1}_{n_2}^\top + v\,\mathbf{I}_{n_2}\right)$$

Optimal design

Exact design:

$$\xi := \begin{pmatrix} \text{G1} & \text{G2} \\ w_1 & w_2 \end{pmatrix}$$

$$w_1 = \frac{n_1}{N} \quad \& \quad w_2 = \frac{n_2}{N}$$

Approximate design:

$$\xi := \begin{pmatrix} \text{G1} & \text{G2} \\ w & 1-w \end{pmatrix}; \quad 0 \leq w \leq 1$$

Search for optimal weight w^* to minimize

a) variance of $\hat{\alpha}_0$

b) MSE matrix of $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_1, ..., \hat{\alpha}_N)^\top$

A-optimal design for estimation of population parameter

A-criterion for estimation of population parameter α_0

$$\Phi_{A,\alpha_0} := \mathrm{var}\left(\hat{\alpha}_0\right)$$

$$\text{for} \quad \xi = \begin{pmatrix} \text{G1} & \text{G2} \\ w & 1-w \end{pmatrix}$$

$$\Phi_{A,\alpha_0}(w) = \frac{K(v+u)+1}{w} + \frac{Ku+1}{1-w}$$

Optimal weight:

$$w^*_{A,\alpha_0} = \frac{1}{1+\sqrt{\frac{Ku+1}{K(v+u)+1}}}$$

D-optimal design for estimation of population parameter

D-criterion for estimation of population parameter α_0

$$\Phi_{D,\alpha_0} := \log\left(\operatorname{var}\left(\hat{\alpha}_0\right)\right)$$

$$\text{for} \quad \xi = \begin{pmatrix} \text{G1} & \text{G2} \\ w & 1-w \end{pmatrix}$$

$$\Phi_{D,\alpha_0}(w) = \log\left(\frac{K(v+u)+1}{w} + \frac{Ku+1}{1-w}\right)$$

Optimal weight:

$$w^*_{D,\alpha_0} = w^*_{A,\alpha_0}$$

A-optimal design for prediction of individual parameters

A-criterion for prediction of individual parameters $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)^\top$:

$$\Phi_{A,\alpha} := \mathrm{tr}\left(\mathrm{Cov}\left(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\right)\right)$$

for $\xi = \begin{pmatrix} \text{G1} & \text{G2} \\ w & 1-w \end{pmatrix}$

$$\Phi_{A,\alpha}(w) = \frac{K(v+u)+1}{w} + \frac{Ku+1}{1-w} - \frac{K^2 v^2 N w}{K(v+u)+1}$$

No formula for $w^*_{A,\alpha}$ $\Rightarrow$ Numerical example

D-optimal design for prediction of individual parameters

D-criterion for prediction of individual parameters $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)^\top$:

$$\Phi_{D,\alpha} := \log\left(\det\left(\mathrm{Cov}\left(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\right)\right)\right)$$

for $\xi = \begin{pmatrix} \text{G1} & \text{G2} \\ w & 1-w \end{pmatrix}$

$$\Phi_{D,\alpha}(w) = N\, w \log\left(\frac{Ku+1}{K(v+u)+1}\right) - \log\left(w(1-w)\right)$$

Optimal weight:

$$w^*_{D,\alpha} = \frac{1}{2} + \frac{1}{Nt} + \sqrt{\frac{1}{4} + \frac{1}{(Nt)^2}}, \qquad t = \log\left(\frac{Ku+1}{K(v+u)+1}\right)$$

A-optimal design for prediction of individual parameters

A-optimal weight $w^*_{A,\alpha}$

of treatment group G1

for prediction of individual parameters

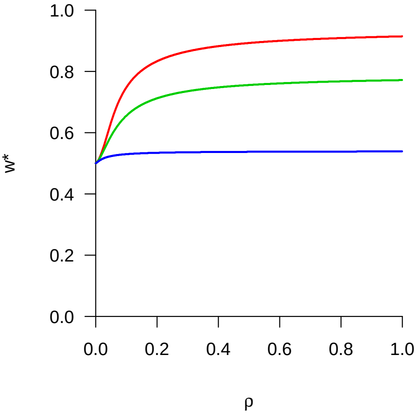
$$q = \frac{v}{u}$$

- $q = 2$
- $q = 0.5$
- $q = 0.1$

$\rho = v/(1+v)$

$N = 100$

$K = 5$



Efficiency of OD in fixed effects model

Efficiency of *A*-optimal design

in fixed effects model $w_A^* = 0.5$

$$eff = \frac{\Phi_{A,\alpha}(w_{A,\alpha}^*)}{\Phi_{A,\alpha}(w_A^*)}$$

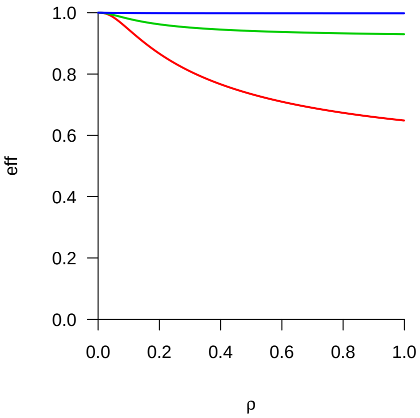
$$q = \frac{v}{u}$$

- $q = 2$
- $q = 0.5$
- $q = 0.1$

$\rho = v/(1+v)$

$N = 100$

$K = 5$



Efficiency of OD for estimation of population parameters

Efficiency of A-optimal design

for estimation w^*_{A,α_0}

$$eff = \frac{\Phi_{A,\alpha}(w^*_{A,\alpha})}{\Phi_{A,\alpha}(w^*_{A,\alpha_0})}$$

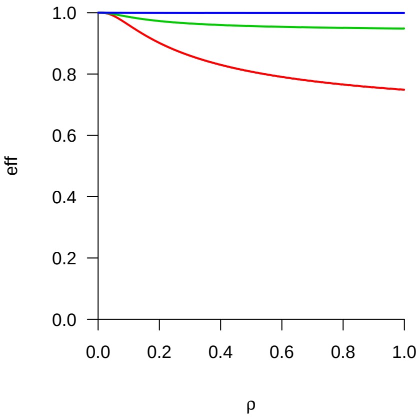
$$q = \frac{v}{u}$$

- $q = 2$
- $q = 0.5$
- $q = 0.1$

$\rho = v/(1+v)$

$N = 100$

$K = 5$



D-optimal design for prediction of individual parameters

D-optimal weight w^*

of treatment group G1

for prediction of individual parameters

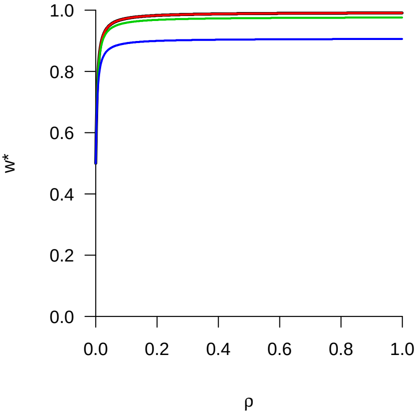
$q = \frac{v}{u}$

- $q = 2$
- $q = 0.5$
- $q = 0.1$

$\rho = v/(1 + v)$

$N = 100$

$K = 5$



Efficiency of OD in fixed effects model

Efficiency of D-optimal design

in fixed effects model $w_D^* = 0.5$

$$eff = \left(\frac{\exp(\Phi_{D,\alpha}(w_{D,\alpha}^*))}{\exp(\Phi_{D,\alpha}(w_D^*))} \right)^{\frac{1}{N}}$$

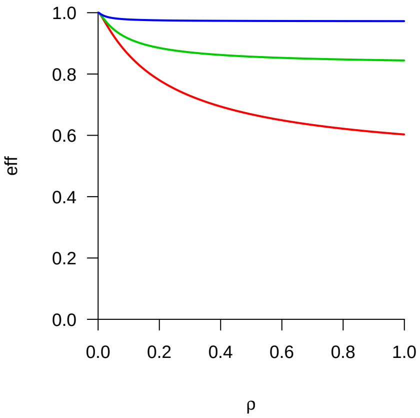
$$q = \frac{v}{u}$$

- $q = 2$
- $q = 0.5$
- $q = 0.1$

$\rho = v/(1+v)$

$N = 100$

$K = 5$



Efficiency of OD for estimation of population parameters

Efficiency of D-optimal design

for estimation w^*_{D,α_0}

$$eff = \left(\frac{\exp(\Phi_{D,\alpha}(w^*_{D,\alpha}))}{\exp(\Phi_{D,\alpha}(w^*_{D,\alpha_0}))} \right)^{\frac{1}{N}}$$

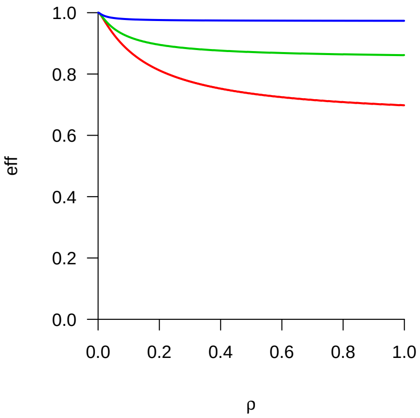
$q = \dfrac{v}{u}$

- $q = 2$
- $q = 0.5$
- $q = 0.1$

$\rho = v/(1+v)$

$N = 100$

$K = 5$



Directions for future research

- Other design criteria
- More than two groups
- Unknown variance parameters
- OD for prediction of other linear aspects
- ...

Thank you for your attention!