Center-Adjusted Inference for a Nonparametric Bayesian Random Effect Distribution

Yisheng Li, Peter Müller and Xihong Lin
M.D. Anderson Cancer Center and Harvard U.

Slide 1

Outline

1. Intro, notation and DP priors for random effects
2. Identifiability issue – everybody knows it, but nobody fixed it
3. Adjusted inference
4. Posterior moments of the (random) moments μ_G and Cov_G of G
5. MCMC implementation
6. Default priors
7. Simulation study
8. Data example
9. Conclusion

1 Introduction

Slide 2

Intro

Mixed effects model:

$$y_{ij} = \beta_0 + \beta_1 s_{ij} + b_{0i} + b_{1i} s_{ij} + \epsilon_{ij}$$

log-PSA for i-th patient at time s_{ij}.
Fixed effects (β_0, β_1) and random effects $b_i = (b_{0i}, b_{1i})$

NP Bayes random effects distribution: heterogeneous patient population

$$b_i \sim G \text{ and } G \sim \text{DP}(G_0, M).$$

with $G_0 = N(0, D)$ and $p(D)$.

Notation:

- Fixed effects (β_0, β_1) are *paired* with b_i. (will argue: adjust for – random – mean of G)
- Variance components: elements of D (will argue: adjust for – random – variance-covariance matrix of G).

2 Identifiability

Slide 3

Identifiability

Random moments of G: random moments of the random prob measure G.

$$\mu_G = \int b \, dG(b) \text{ and } \text{Cov}_G = \int (b - \mu_G)(b - \mu_G)^t dG(b)$$

Of course, $\mu_G \neq 0$ a.s. and $\text{Cov}_G \neq D$ (even true for prior expectation).

Fixed effects: Shifting $\mu_G \to \mu_G + c$ (if doable) and $\beta \to \beta - c$ leaves likelihood unchanged.
Nothing new, but nobody addresses it.

3 Center-Adjusted Inference

Slide 4

Adjusted Inference

Adjusted inference: Report inference on $\beta + \mu_G$ as fixed effect.
Can be implemented by simple post-processing of MCMC output.

Variance components: Similarly, report $p(\text{Cov}_G \mid \textit{data})$ instead of $p(D \mid \textit{data})$.
Simple post-processing of MCMC output.

GLMM: Remains valid for any GLMM with linear predictor

$$\eta = \beta' x_i + b_i' z_i = \beta_F' x_{F,i} + \beta_R' x_{R,i} + b_i' x_{R,i}$$

with some paired (matching columns in the design matrix) fixed and random effects.

Slide 5

Centered Model

Centered Model: Move β into the location parameter of the base measure

$$\eta = \beta_F' x_{F,i} + b_i' x_{R,i}$$
$$b_i \sim \text{DP}(G_0, M) \text{ and } G_0 = N(\beta_R, D)$$

Adjusted inference – fixed effects: Report $E(\mu_G \mid \textit{data})$ and $\text{Var}(\mu_G \mid \textit{data})$ for fixed effects.
Construct C.I. by matching normal moments.

Adjusted inference – variance components: Report $E(\mathrm{Cov}_G \mid data)$ and $\mathrm{Var}(\mathrm{Cov}_G \mid data)$.

Implementation: The required moments of μ_G and Cov_G are evaluated using existing MCMC output → see second part.

Slide 6

Alternatives

Approximation based on sampling G: Gelfand & Mukhopadhyay (1995) and Gelfand & Kottas (2002)

- sample (approx) G conditional on imputed parameters at each iteration
- compute μ_G based on sampled G
- similar for higher moments

Restrict G: restrict G to zero median. Easy with PT prior.

Results on $p(\mu_G)$: Hjort and Ongaro (2005), Lijoi & Regazzini (2004), Epifani, Guglilmi & Melilli (2006) (including Cov_G).

4 Posterior Moments of μ_G and Cov_G

Slide 7

Posterior Moments of μ_G

Let

- $G_\star = \{M \cdot N(\beta_R, D) + m \cdot \overline{\delta_{b_i}}\}/(M+m)$ with $\overline{\delta_{b_i}} = \sum_{i=1}^m \delta_{b_i}/m$
- $\mu_{G_\star} = \int b dG_\star(b) = (M\beta_R + \sum_{i=1}^m b_i)/(M+m)$
- $\mathrm{Cov}_{G_\star} = \int (b-\mu_{G_\star})(b-\mu_{G_\star})^t dG_\star(b)$

Then

Proposition 1. (i) $E(\mu_G \mid y) = E(\mu_{G_\star} \mid y)$;
(ii) $\mathrm{Cov}(\mu_G \mid y) = E\left(\frac{\mathrm{Cov}_{G_\star}}{m+M+1} \mid y\right) + \mathrm{Cov}(\mu_{G_\star} \mid y)$.

Slide 8

Posterior Moments of Cov_G

Proposition 2.

1. $E(\mathrm{Cov}_G \mid y) = E\left(\frac{m+M}{m+M+1} \cdot \mathrm{Cov}_{G_\star} \mid y\right)$
2. $\mathrm{Cov}(\mathrm{Cov}_{G,i_1j_1}, \mathrm{Cov}_{G,i_2j_2} \mid y) = \ldots$ (closed form expression – and you don't want to see it now!) $\mathrm{Var}(\mathrm{Cov}_{G,ij} \mid y) = \ldots$ as special cases.

5 MCMC Implementation

Slide 9

Posterior Inference

Gibbs or Metropolis-within-Gibbs algorithm:

1. Direct (when conjugate) or adaptive rejection sampling for β_F, β_R and b_i
2. Direct sampling or a Metropolis step for VC (D, σ^2)

Inference:

- For $z_{ij}^t b_i$, i.e., random effect as slope for z_{ij}, report
    1. $E(\mu_G \mid y)$ as point estimate for the fixed effect
    2. $\mathrm{Cov}(\mu_G \mid y)$ as posterior covariance
- Report as inference for random effect covariance:
    1. $E(\mathrm{Cov}_{G,ij} \mid y)$ as point estimates
    2. $\mathrm{Var}(\mathrm{Cov}_{G,ij} \mid y)$ for posterior variances

6 Default Priors and Posterior Propriety

Slide 10

Prior Specification and Posterior Propriety

Proposition 3. $\pi(\beta_F, \beta_R, \sigma^2) = 1/\sigma^2$, $\pi(D)$ proper ⇒ proper posterior

(same w/o σ^2 for non-normal outcome under appropriate conditions: DP GLMM)

Examples of $\pi(D)$: uniform shrinkage prior (Natarajan and Kass, 2000; Daniels, 1999), IWP

7 Simulation Studies

Slide 11

Simulation Study: Normal Outcome

- 200 datasets
- 50 subjects
- 10 obs per subject
- intercept + slope (fixed and random)
- random effects: normal mixture
- $\beta_0 = \beta_1 = 1$, $x_{ij} = j + .025i - 5$, $\sigma^2 = 1$

$$Y_{ij} = \beta_0 + b_i^{(1)} + \left(\beta_1 + b_i^{(2)}\right) x_{ij} + \epsilon_{ij}$$

Simulation Study: Normal Outcome (ctd) – Summary

Comparison btw Unadj and Adj (comparable for σ^2):

- Bias: 3-15 times
- MSE: > 2 times
- CIL: > 2 times
- CP: comparable

Comparison btw IWP and USP:

- CIL for VC (comparable for σ^2): > 1.25 times
- Bias, MSE, CP: comparable

Par	$\pi(\boldsymbol{D})$	Center-adjusted				Conventional			
		Bias	MSE	CIL	CP	Bias	MSE	CIL	CP
β_0	IWP	.04	.04	.85	.93	.16	.08	2.88	.99
	USP	.03	.04	.81	.93	.15	.09	1.89	.97
β_1	IWP	-.04	.04	.84	.93	-.15	.08	1.88	.80
	USP	-.03	.04	.80	.95	-.15	.09	1.60	.85
σ^2	IWP	.04	.01	.29	.94	.04	.01	.29	.94
	USP	.04	.01	.29	.92	.03	.01	.29	.94
σ_{11}	IWP	.01	.10	1.62	.99	.15	.33	2.44	.99
	USP	-.07	.11	1.29	.93	-.20	.31	1.46	.45
σ_{22}	IWP	.01	.08	1.53	1.00	.16	.30	4.31	.96
	USP	-.06	.09	1.19	.96	-.20	.31	2.49	.98
σ_{12}	IWP	-.01	.09	1.54	.99	-.15	.30	4.16	1.00
	USP	.06	.09	1.20	.94	.20	.31	2.44	.99

$M \sim G(2.5, .5)$

Simulation Study: Binary Outcome

- 200 datasets
- 100 subjects
- 10 obs per subject
- bivariate random effects

$$Y_{ij} \mid b_i \sim Bernoulli(p_{ij}),$$

$$logit(p_{ij}) = \beta_0 + b_i^{(1)} + \left(\beta_1 + b_i^{(2)}\right) \cdot x_{ij}.$$

Simulation Study: Binary Outcome (ctd) – Summary

Comparison btw Unadj and Adj:

- Bias: 1-9 times
- MSE: 1-3 times
- CIL: 1.3-2.5 times
- CP: ≥97% vs 92-97%

Comparison btw IWP and USP:

- Fixed effects: Bias: 3% vs 6-7%, MSE: comparable, CIL: ≈ 1.1 times, CP: 97% vs 93-94%
- VC:
  - Bias: IWP: 22-34% vs USP: 2-3%
  - MSE: 2-4 times
  - CIL: ≈ 1.5 times
  - CP: ≥ 99% vs 92-97%

Simulation study: binary outcome (ctd)
$M = 5$

Par	$\pi(\boldsymbol{D})$	Center-adjusted				Conventional			
		Bias	MSE	CIL	CP	Bias	MSE	CIL	CP
β_0	IWP	.03	.06	1.01	.97	.04	.13	2.68	1.00
	USP	.07	.05	.91	.94	.24	.14	2.32	1.00
β_1	IWP	.03	.07	1.16	.97	-.03	.13	2.70	1.00
	USP	-.06	.06	.96	.93	-.25	.14	2.19	1.00
σ_{11}	IWP	.22	1.26	4.77	.99	.51	3.83	10.11	1.00
	USP	.02	.57	3.31	.97	-.19	.58	4.54	.99
σ_{22}	IWP	.34	2.04	6.26	.99	.50	3.80	10.34	1.00
	USP	-.02	.49	3.67	.97	-.29	.77	4.18	.97
σ_{12}	IWP	-.27	1.32	5.26	.99	.50	3.33	9.80	1.00
	USP	.03	.39	3.13	.92	-.27	.67	4.14	.99

8 Data Example

Slide 17

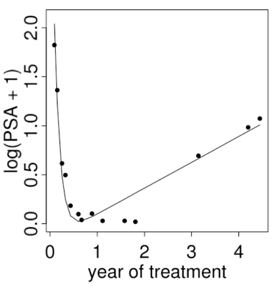
PSA Study

- Outcome $y = \log(\text{PSA} + 1)$
- Two trt arms: CH (137 pts) vs AA (149 pts)
- 1-65 obs per subject
- s_{vij}: time post trt init
- $v = 0$ (CH), 1 (AA)

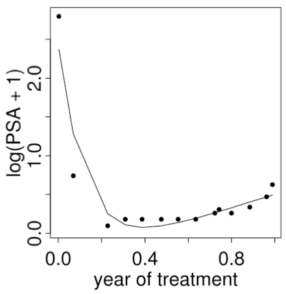
$$y_{vij} = \mu_0+\theta_{0vi}+(\theta_{1vi}+v\cdot d_g)s_{vij}+(\theta_{2vi}+v\cdot d_\eta)\left(e^{-\phi_v s_{vij}}-1\right)+\epsilon_{vij}$$

Slide 18

PSA Study (ctd) – Individual PSA Profiles



(ID: 38)



(ID: 168)

Slide 19

PSA Study (ctd)

Posterior summary

Arm	Par	Adj	PM	PSD	95% CI
		Init drop rate			
CH	ϕ_0	Adj/Un	8.44	.21	(8.04, 8.87)
AA	ϕ_1	Adj/Un	8.03	.20	(7.63, 8.44)
		Increase per year			
CH	μ_{g_1}	Adj	.63	.08	(.49, .78)
	β_1	Unadj	.70	.25	(.24, 1.24)
AA	$\mu_{g_1} + d_g$	Adj	.62	.08	(.47, .77)
	$\beta_1 + d_g$	Un	.69	.25	(.21, 1.22)
		Init drop			
CH	μ_{g_2}	Adj	3.32	.14	(3.04, 3.59)
	β_2	Un	4.33	.48	(3.37, 5.28)
AA	$\mu_{g_2} + d_\eta$	Adj	3.17	.14	(2.89, 3.44)
	$\beta_2 + d_\eta$	Un	4.18	.48	(3.23, 5.14)

Slide 20

PSA Study (ctd)

Posterior summary (ctd)

Par	Adj	PM	PSD	95% CI
			VC	
σ_{11}	Adj	1.17	.23	(.78, 1.68)
	Un	1.76	.68	(.89, 3.54)
σ_{22}	Adj	4.76	.51	(3.84, 5.84)
	Un	7.82	2.05	(4.65, 12.56)
σ_{12}	Adj	.35	.17	(.01, .68)
	Un	-.22	.60	(-1.70, 1.14)

9 Summary

Slide 21

Summary

Conventional inference: bias in small samples, poor inference

Center-adjusted inference

- Unbiased, satisfactory inference
- Applicable to ANY DP mixture model

R code for post-processing: `http://odin.mdacc.tmc.edu/~yishengli/DPPP.R`

Paper: `http://odin.mdacc.tmc.edu/~yishengli/dp.pdf`