Bayesian Inference for skew-normal linear mixed models with covariates measurements errors.

Oludare Ariyo[1,2] Emmanuel Lesaffre[2]

[1]L-BioStat, Katholieke Universiteit, Leuven Belgium

[2]Department of Statistics
Federal University of Agriculture, Abeokuta, Nigeria

ISCB 2017 Students Day

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Key Massages/ Questions

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Key Massages/ Questions

- Key Massages

Interuniversity Institute for Biostatistics
and statistical Bioinformatics

Key Massages/ Questions

- Key Massages
  - Models assumptions is critical in data analysis

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Key Massages/ Questions

- Key Massages
  - Models assumptions is critical in data analysis
  - Quality research are not straightforward, they require extra effort

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Key Massages/ Questions

- Key Massages
  - Models assumptions is critical in data analysis
  - Quality research are not straightforward, they require extra effort
  - Simulation studies setting: Important issues to deal with

Key Massages/ Questions

- Key Massages
  - Models assumptions is critical in data analysis
  - Quality research are not straightforward, they require extra effort
  - Simulation studies setting: Important issues to deal with
- Questions

Key Massages/ Questions

- Key Massages
  - Models assumptions is critical in data analysis
  - Quality research are not straightforward, they require extra effort
  - Simulation studies setting: Important issues to deal with
- Questions
  - What to do when your model cannot be implement in available software

UHASSELT I-BioStat KU LEUVEN
Interuniversity Institute for Biostatistics and statistical Bioinformatics

Key Massages/ Questions

- Key Massages
  - Models assumptions is critical in data analysis
  - Quality research are not straightforward, they require extra effort
  - Simulation studies setting: Important issues to deal with
- Questions
  - What to do when your model cannot be implement in available software
  - Computational difficulties !, when model is not convergence or unnecessarily slow.

Key Massages/ Questions

- Key Massages
  - Models assumptions is critical in data analysis
  - Quality research are not straightforward, they require extra effort
  - Simulation studies setting: Important issues to deal with
- Questions
  - What to do when your model cannot be implement in available software
  - Computational difficulties !, when model is not convergence or unnecessarily slow.
  - What your simulation results 'contradict' real life data analysis

UHASSELT I-BioStat KU LEUVEN
Interuniversity Institute for Biostatistics and statistical Bioinformatics

Key Massages/ Questions

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Key Massages/ Questions

- Background

Key Massages/ Questions

- Background
  - B.Sc Mathematics

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Key Massages/ Questions

- Background
  - B.Sc Mathematics
  - M.Sc Statistic (Mathematics)

Interuniversity Institute for Biostatistics
and statistical Bioinformatics

Key Massages/ Questions

- Background
    - B.Sc Mathematics
    - M.Sc Statistic (Mathematics)
- PhD Projects

UHASSELT I-BioStat KU LEUVEN
Interuniversity Institute for Biostatistics and statistical Bioinformatics

Key Massages/ Questions

- Background
  - B.Sc Mathematics
  - M.Sc Statistic (Mathematics)
- PhD Projects
  - Longitudinal Data Analysis, Model selection in Bayesian Paradigm

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Key Massages/ Questions

- Background
  - B.Sc Mathematics
  - M.Sc Statistic (Mathematics)
- PhD Projects
  - Longitudinal Data Analysis, Model selection in Bayesian Paradigm
  - examine and develop statistical tools in a Bayesian context the appropriateness of diagnostic tools for overall predictive performance of an assumed mixed model

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Key Massages/ Questions

- Background
  - B.Sc Mathematics
  - M.Sc Statistic (Mathematics)
- PhD Projects
  - Longitudinal Data Analysis, Model selection in Bayesian Paradigm
  - examine and develop statistical tools in a Bayesian context the appropriateness of diagnostic tools for overall predictive performance of an assumed mixed model
  - to find the reason for specific model deviations such as the presence of outliers and influential observations

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Linear mixed model (LMM)

$$\begin{cases} Y_i = X_i\beta + Z_i b_i + \epsilon_i \\ b_i \sim N_q(0, G) \\ \epsilon_i \sim N_{n_i}(0, \Sigma_i) \\ b_1, \ldots, b_n, \epsilon_1, \ldots, \epsilon_n \text{ independent,} \end{cases} \tag{1}$$

- where
  - G is a $(q \times q)$ covariance matrix.
  - Σ_i is a $(n_i \times n_i)$ depends on i only

UHASSELT I-BioStat KU LEUVEN
Interuniversity Institute for Biostatistics and statistical Bioinformatics

Linear mixed model (LMM)

$$\begin{cases} Y_i = X_i\beta + Z_i b_i + \epsilon_i \\ b_i \sim N_q(0, G) \\ \epsilon_i \sim N_{n_i}(0, \Sigma_i) \\ b_1, \ldots, b_n, \epsilon_1, \ldots, \epsilon_n \text{ independent,} \end{cases} \tag{1}$$

- where
  - G is a $(q \times q)$ covariance matrix.
  - Σ_i is a $(n_i \times n_i)$ depends on i only
- Assumes normality for the random error and random effects

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Linear mixed model (LMM)

$$\begin{cases} Y_i = X_i\beta + Z_ib_i + \epsilon_i \\ b_i \sim N_q(0, G) \\ \epsilon_i \sim N_{n_i}(0, \Sigma_i) \\ b_1, \ldots, b_n, \epsilon_1, \ldots, \epsilon_n \text{ independent,} \end{cases} \quad (1)$$

- where
  - G is a $(q \times q)$ covariance matrix.
  - Σ_i is a $(n_i \times n_i)$ depends on i only
- Assumes normality for the random error and random effects
- Assumptions? (Verbeke and Lesaffre, 1997; Lachos et al., 2010)

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Linear mixed model (LMM)

$$\begin{cases} Y_i = X_i\beta + Z_ib_i + \epsilon_i \\ b_i \sim N_q(0, G) \\ \epsilon_i \sim N_{n_i}(0, \Sigma_i) \\ b_1, \ldots, b_n, \epsilon_1, \ldots, \epsilon_n \text{ independent,} \end{cases} \quad (1)$$

- where
  - G is a $(q \times q)$ covariance matrix.
  - Σ_i is a $(n_i \times n_i)$ depends on i only
- Assumes normality for the random error and random effects
- Assumptions? (Verbeke and Lesaffre, 1997; Lachos et al., 2010) when the data present skewness

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Linear mixed model (LMM)

$$\begin{cases} Y_i = X_i\beta + Z_i b_i + \epsilon_i \\ b_i \sim N_q(0, G) \\ \epsilon_i \sim N_{n_i}(0, \Sigma_i) \\ b_1, \ldots, b_n, \epsilon_1, \ldots, \epsilon_n \text{ independent,} \end{cases} \quad (1)$$

- where
  - G is a $(q \times q)$ covariance matrix.
  - Σ_i is a $(n_i \times n_i)$ depends on i only
- Assumes normality for the random error and random effects
- Assumptions? (Verbeke and Lesaffre, 1997; Lachos et al., 2010) when the data present skewness
- ... We consider some extensions in Bayesian paradigm

Bayesian model selection . . .

. . . (extensions of LMM) models,

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Bayesian model selection . . .

. . . (extensions of LMM) models,

- DIC (Spiegelhalter et al., 2002) measure the predictive ability of the fitted model to future data.

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Bayesian model selection . . .

. . . (extensions of LMM) models,

- DIC (Spiegelhalter et al., 2002) measure the predictive ability of the fitted model to future data. Theoretical Justification?

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Bayesian model selection . . .

. . . (extensions of LMM) models,

- DIC (Spiegelhalter et al., 2002) measure the predictive ability of the fitted model to future data. Theoretical Justification?
- Celeux et al. (2006) discussed different definitions for DIC, including

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Bayesian model selection . . .

. . . (extensions of LMM) models,

- DIC (Spiegelhalter et al., 2002) measure the predictive ability of the fitted model to future data. Theoretical Justification?

- Celeux et al. (2006) discussed different definitions for DIC, including
  - conditional DIC (cDIC) →calculated based on the conditional likelihood $p(y|\phi, \mu)$

Bayesian model selection . . .

. . . (extensions of LMM) models,

- DIC (Spiegelhalter et al., 2002) measure the predictive ability of the fitted model to future data. Theoretical Justification?
- Celeux et al. (2006) discussed different definitions for DIC, including
  - conditional DIC (cDIC) →calculated based on the conditional likelihood $p(y|\phi, \mu)$
  - marginal DIC (mDIC) →$p(y|\phi) = \int_{\mu} p(y|\phi, \mu)p(\mu|\phi)d\mu$
    - ϕ →vector of parameters
    - μ →latent variables (random effect)

Bayesian model selection . . .

. . . (extensions of LMM) models,

Bayesian model selection . . .

. . . (extensions of LMM) models,

- It has been shown via simulation that mDIC outperforms cDIC since the latter tends to select the over-fitted model (Chan and Grant, 2016).

Bayesian model selection . . .

. . . (extensions of LMM) models,

- It has been shown via simulation that mDIC outperforms cDIC since the latter tends to select the over-fitted model (Chan and Grant, 2016).
- Computational difficulties are the major drawback for mDIC

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Bayesian model selection . . .

. . . (extensions of LMM) models,

- It has been shown via simulation that mDIC outperforms cDIC since the latter tends to select the over-fitted model (Chan and Grant, 2016).
- Computational difficulties are the major drawback for mDIC
- mDIC computation via importance sampling.

Bayesian model selection . . .

. . . (extensions of LMM) models,

- It has been shown via simulation that mDIC outperforms cDIC since the latter tends to select the over-fitted model (Chan and Grant, 2016).
- Computational difficulties are the major drawback for mDIC
- mDIC computation via importance sampling.
- We extend the importance sampling algorithms for computation of mDIC to the model with skew-normal latent variables

UHASSELT I-BioStat KU LEUVEN
Interuniversity Institute for Biostatistics and statistical Bioinformatics

Simulation studies

- $$y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 g_i + b_i + \epsilon_{ij}, \quad (2)$$

- $n = 184$ with g_i=0 if $i \leq 92$ and g_i=1 if $i > 92$,
- $\beta_0 = 4$, $\beta_1 = 1$ and $\beta_2 = 2$.

- generate 100 Monte Carlo data from equation (2) using the R software jointly with *rjags* with the following specifications
  - $\beta_0, \beta_1, \beta_2 \sim N_1\left(0, 10^2\right)$,
  - $\sigma_\epsilon^2, \sigma_e^2 \sim IG\left(0.01, 0.01\right)$,
  - $\delta_b \sim N_1\left(0, 10^2\right) I\left\{\delta_b > 0\right\}$.

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Simulation results

Table: The results of Mento carlo based on 100 generated data sets, $N_1(0,4)$ distribution for the random effects

Parameter	Real	MC Mean	MC SD	MC Median	5% th.q	95% th.q
(a) Normal Scenario						
β_0	4	4.0597	0.0031	4.0599	4.0536	4.0669
β_1	2	1.7586	0.0104	1.7582	1.7403	1.7836
β_2	1	1.0520	0.0146	1.0505	1.0109	1.0849
σ_ε^2	0.25	0.2477	0.0014	0.2476	0.2450	0.2507
σ_b^2		4.0278	0.0011	4.0278	4.0260	4.0303
(b) Skew-Normal Scenario						
β_0	4	5.6095	0.2892	5.6230	5.1030	6.1008
β_1	2	2.1318	0.0111	2.1324	2.4895	2.1707
β_2	1	0.9014	0.0166	0.9063	0.5765	1.1862
σ_ε^2	0.25	0.2453	0.0174	0.2436	0.2275	0.2579
σ_b^2	-	3.8217	0.1521	3.8270	3.5362	4.0762
δ_b	-	1.0491	0.6013	1.0491	1.0464	1.0516

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Simulation results

Table: The results of Mento carlo based on 100 generated data sets, $N_1(0,4)$ distribution for the random effects

Parameter	Real	MC Mean	MC SD	MC Median	5% th.q	95% th.q
(a) Normal Scenario						
β_0	4	4.0597	0.0031	4.0599	4.0536	4.0669
β_1	2	1.7586	0.0104	1.7582	1.7403	1.7836
β_2	1	1.0520	0.0146	1.0505	1.0109	1.0849
σ_ξ^2	0.25	0.2477	0.0014	0.2476	0.2450	0.2507
σ_b^2		4.0278	0.0011	4.0278	4.0260	4.0303
(b) Skew-Normal Scenario						
β_0	4	5.6095	0.2892	5.6230	5.1030	6.1008
β_1	2	2.1318	0.0111	2.1324	2.4895	2.1707
β_2	1	0.9014	0.0166	0.9063	0.5765	1.1862
σ_ξ^2	0.25	0.2453	0.0174	0.2436	0.2275	0.2579
σ_b^2	-	3.8217	0.1521	3.8270	3.5362	4.0762
δ_b	-	1.0491	0.6013	1.0491	1.0464	1.0516

- For most part, parameter estimates are unbiased

UHASSELT I-BioStat KU LEUVEN
Interuniversity Institute for Biostatistics and statistical Bioinformatics

Simulation results

Table: The results of Mento carlo based on 100 generated data sets, $N_1(0,4)$ distribution for the random effects

Parameter	Real	MC Mean	MC SD	MC Median	5% th.q	95% th.q
(a) Normal Scenario						
β_0	4	4.0597	0.0031	4.0599	4.0536	4.0669
β_1	2	1.7586	0.0104	1.7582	1.7403	1.7836
β_2	1	1.0520	0.0146	1.0505	1.0109	1.0849
σ_ϵ^2	0.25	0.2477	0.0014	0.2476	0.2450	0.2507
σ_b^2		4.0278	0.0011	4.0278	4.0260	4.0303
(b) Skew-Normal Scenario						
β_0	4	5.6095	0.2892	5.6230	5.1030	6.1008
β_1	2	2.1318	0.0111	2.1324	2.4895	2.1707
β_2	1	0.9014	0.0166	0.9063	0.5765	1.1862
σ_ϵ^2	0.25	0.2453	0.0174	0.2436	0.2275	0.2579
σ_b^2	-	3.8217	0.1521	3.8270	3.5362	4.0762
δ_b	-	1.0491	0.6013	1.0491	1.0464	1.0516

- For most part, parameter estimates are unbiased
- posterior distributions are fairly symmetric

Simulation results

Table: The results of Mento carlo based on 100 generated data sets, $N_1(0,4)$ distribution for the random effects

Parameter	Real	MC Mean	MC SD	MC Median	5% th.q	95% th.q
(a) Normal Scenario						
β_0	4	4.0597	0.0031	4.0599	4.0536	4.0669
β_1	2	1.7586	0.0104	1.7582	1.7403	1.7836
β_2	1	1.0520	0.0146	1.0505	1.0109	1.0849
σ_ϵ^2	0.25	0.2477	0.0014	0.2476	0.2450	0.2507
σ_b^2		4.0278	0.0011	4.0278	4.0260	4.0303
(b) Skew-Normal Scenario						
β_0	4	5.6095	0.2892	5.6230	5.1030	6.1008
β_1	2	2.1318	0.0111	2.1324	2.4895	2.1707
β_2	1	0.9014	0.0166	0.9063	0.5765	1.1862
σ_ϵ^2	0.25	0.2453	0.0174	0.2436	0.2275	0.2579
σ_b^2	-	3.8217	0.1521	3.8270	3.5362	4.0762
δ_b	-	1.0491	0.6013	1.0491	1.0464	1.0516

- For most part, parameter estimates are unbiased
- posterior distributions are fairly symmetric
- penalty for estimating the random effects density when the normal assumption holds is minimal

UHASSELT I-BioStat KU LEUVEN
Interuniversity Institute for Biostatistics and statistical Bioinformatics

Simulation results

Table: The results of Mento carlo based on 100 generated data sets, $N_1(0, 4)$ distribution for the random effects

Parameter	Real	MC Mean	MC SD	MC Median	5% th.q	95% th.q
(a) Normal Scenario						
β_0	4	4.0597	0.0031	4.0599	4.0536	4.0669
β_1	2	1.7586	0.0104	1.7582	1.7403	1.7836
β_2	1	1.0520	0.0146	1.0505	1.0109	1.0849
σ_ξ^2	0.25	0.2477	0.0014	0.2476	0.2450	0.2507
σ_b^2		4.0278	0.0011	4.0278	4.0260	4.0303
(b) Skew-Normal Scenario						
β_0	4	5.6095	0.2892	5.6230	5.1030	6.1008
β_1	2	2.1318	0.0111	2.1324	2.4895	2.1707
β_2	1	0.9014	0.0166	0.9063	0.5765	1.1862
σ_ξ^2	0.25	0.2453	0.0174	0.2436	0.2275	0.2579
σ_b^2	-	3.8217	0.1521	3.8270	3.5362	4.0762
δ_b	-	1.0491	0.6013	1.0491	1.0464	1.0516

- For most part, parameter estimates are unbiased
- posterior distributions are fairly symmetric
- penalty for estimating the random effects density when the normal assumption holds is minimal
- The mDIC and cDIC criteria correctly selected the normal distribution 98% and 96% respectively

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Simulation results

Table: The results of Mento Carlo based on 100 generated data sets, Gamma$(4, 1)$ distribution for the random effects

Parameter	Real	MC Mean	MC SD	MC Median	5% th.q	95% th.q
(a) Normal Scenario						
β_0	4	3.8432	0.3012	3.8431	3.5237	4.0124
β_1	2	2.0516	0.0122	2.0517	2.0469	2.0558
β_2	1	1.0931	0.6114	1.0915	1.0731	1.1132
σ_ϵ^2	0.25	0.2516	0.0011	0.2515	0.2498	0.2538
σ_b^2		4.3972	1.1997	4.3841	4.1901	4.6641
(b) Skew-Normal Scenario						
β_0	4	1.5755	0.0016	1.5776	1.5425	1.6011
β_1	2	2.0879	0.0002	2.0879	2.0833	2.0922
β_2	1	0.8964	0.0071	0.8962	0.8849	0.8105
σ_ϵ^2	0.25	0.2482	0.0072	0.2479	0.2448	0.2811
σ_b^2		1.0664	0.0014	1.0665	1.0641	1.0692
δ_b		3.0132	0.0140	3.0132	0.0132	3.0132

UHASSELT I-BioStat KU LEUVEN
Interuniversity Institute for Biostatistics and statistical Bioinformatics

Simulation results

Table: The results of Mento Carlo based on 100 generated data sets, Gamma$(4,1)$ distribution for the random effects

Parameter	Real	MC Mean	MC SD	MC Median	5% th.q	95% th.q
(a) Normal Scenario						
β_0	4	3.8432	0.3012	3.8431	3.5237	4.0124
β_1	2	2.0516	0.0122	2.0517	2.0469	2.0558
β_2	1	1.0931	0.6114	1.0915	1.0731	1.1132
σ_ϵ^2	0.25	0.2516	0.0011	0.2515	0.2498	0.2538
σ_b^2		4.3972	1.1997	4.3841	4.1901	4.6641
(b) Skew-Normal Scenario						
β_0	4	1.5755	0.0016	1.5776	1.5425	1.6011
β_1	2	2.0879	0.0002	2.0879	2.0833	2.0922
β_2	1	0.8964	0.0071	0.8962	0.8849	0.8105
σ_ϵ^2	0.25	0.2482	0.0072	0.2479	0.2448	0.2811
σ_b^2		1.0664	0.0014	1.0665	1.0641	1.0692
δ_b		3.0132	0.0140	3.0132	0.0132	3.0132

- parameter estimates for most part are unbiased

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Simulation results

Table: The results of Mento Carlo based on 100 generated data sets, Gamma$(4, 1)$ distribution for the random effects

Parameter	Real	MC Mean	MC SD	MC Median	5% th.q	95% th.q
(a) Normal Scenario						
β_0	4	3.8432	0.3012	3.8431	3.5237	4.0124
β_1	2	2.0516	0.0122	2.0517	2.0469	2.0558
β_2	1	1.0931	0.6114	1.0915	1.0731	1.1132
σ_ϵ^2	0.25	0.2516	0.0011	0.2515	0.2498	0.2538
σ_b^2		4.3972	1.1997	4.3841	4.1901	4.6641
(b) Skew-Normal Scenario						
β_0	4	1.5755	0.0016	1.5776	1.5425	1.6011
β_1	2	2.0879	0.0002	2.0879	2.0833	2.0922
β_2	1	0.8964	0.0071	0.8962	0.8849	0.8105
σ_ϵ^2	0.25	0.2482	0.0072	0.2479	0.2448	0.2811
σ_b^2		1.0664	0.0014	1.0665	1.0641	1.0692
δ_b		3.0132	0.0140	3.0132	0.0132	3.0132

- parameter estimates for most part are unbiased
- loss of efficiency of estimate (higher MC SD) on β_1, β_2

Simulation results

Table: The results of Mento Carlo based on 100 generated data sets, Gamma$(4, 1)$ distribution for the random effects

Parameter	Real	MC Mean	MC SD	MC Median	5% th.q	95% th.q
(a) Normal Scenario						
β_0	4	3.8432	0.3012	3.8431	3.5237	4.0124
β_1	2	2.0516	0.0122	2.0517	2.0469	2.0558
β_2	1	1.0931	0.6114	1.0915	1.0731	1.1132
σ_ξ^2	0.25	0.2516	0.0011	0.2515	0.2498	0.2538
σ_b^2		4.3972	1.1997	4.3841	4.1901	4.6641
(b) Skew-Normal Scenario						
β_0	4	1.5755	0.0016	1.5776	1.5425	1.6011
β_1	2	2.0879	0.0002	2.0879	2.0833	2.0922
β_2	1	0.8964	0.0071	0.8962	0.8849	0.8105
σ_ξ^2	0.25	0.2482	0.0072	0.2479	0.2448	0.2811
σ_b^2		1.0664	0.0014	1.0665	1.0641	1.0692
δ_b		3.0132	0.0140	3.0132	0.0132	3.0132

- parameter estimates for most part are unbiased
- loss of efficiency of estimate (higher MC SD) on β_1, β_2
- failure to take appropriate account of the true feature of the random effects leads to less precise inference on what are usually quantiles of interest

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Simulation results

Table: The results of Mento Carlo based on 100 generated data sets, Gamma$(4, 1)$ distribution for the random effects

Parameter	Real	MC Mean	MC SD	MC Median	5% th.q	95% th.q
(a) Normal Scenario						
β_0	4	3.8432	0.3012	3.8431	3.5237	4.0124
β_1	2	2.0516	0.0122	2.0517	2.0469	2.0558
β_2	1	1.0931	0.6114	1.0915	1.0731	1.1132
σ_ξ^2	0.25	0.2516	0.0011	0.2515	0.2498	0.2538
σ_b^2		4.3972	1.1997	4.3841	4.1901	4.6641
(b) Skew-Normal Scenario						
β_0	4	1.5755	0.0016	1.5776	1.5425	1.6011
β_1	2	2.0879	0.0002	2.0879	2.0833	2.0922
β_2	1	0.8964	0.0071	0.8962	0.8849	0.8105
σ_ξ^2	0.25	0.2482	0.0072	0.2479	0.2448	0.2811
σ_b^2		1.0664	0.0014	1.0665	1.0641	1.0692
δ_b		3.0132	0.0140	3.0132	0.0132	3.0132

- parameter estimates for most part are unbiased
- loss of efficiency of estimate (higher MC SD) on β_1, β_2
- failure to take appropriate account of the true feature of the random effects leads to less precise inference on what are usually quantiles of interest
- results are similar to those reported in Hu and Davidian (1998) and Zhang and Davidian (2001) using classical approach and Arellano-Valle et al. (2007) using Bayesian approach.

UHASSELT I-BioStat KU LEUVEN
Interuniversity Institute for Biostatistics and statistical Bioinformatics

Comparing considered models

Table: Comparing competing models using conditional and marginal DIC

Method	Parameters	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
conditional	$\overline{D(\theta)}$	317.56	317.93	319.34	319.91	319.31	321.41
	pD	-569.31	-576.14	-513.21	-582.21	-572.41	-514.71
	$cDIC$	-182.71	-181.91	-184.16	-180.98	-181.31	-188.83
marginal	$D(\bar{\theta})$	390.40	393.03	393.08	398.3	386.40	384.21
	pD	7.48	7.88	10.22	7.47	13.42	12.32
	$mDIC$	415.00	418.90	409.72	416.10	412.10	413.56

Comparing considered models

Table: Comparing competing models using conditional and marginal DIC

Method	Parameters	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
conditional	$\overline{D(\theta)}$	317.56	317.93	319.34	319.91	319.31	321.41
	pD	-569.31	-576.14	-513.21	-582.21	-572.41	-514.71
	$cDIC$	-182.71	-181.91	-184.16	-180.98	-181.31	-188.83
marginal	$D(\bar{\theta})$	390.40	393.03	393.08	398.3	386.40	384.21
	pD	7.48	7.88	10.22	7.47	13.42	12.32
	$mDIC$	415.00	418.90	409.72	416.10	412.10	413.56

- Chan and Grant (2016) showed that cDIC tends to choose over-fitted models while mDIC work better in general

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Comparing considered models

Table: Comparing competing models using conditional and marginal DIC

Method	Parameters	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
conditional	$\overline{D(\theta)}$	317.56	317.93	319.34	319.91	319.31	321.41
	pD	-569.31	-576.14	-513.21	-582.21	-572.41	-514.71
	$cDIC$	-182.71	-181.91	-184.16	-180.98	-181.31	-188.83
marginal	$D(\bar{\theta})$	390.40	393.03	393.08	398.3	386.40	384.21
	pD	7.48	7.88	10.22	7.47	13.42	12.32
	$mDIC$	415.00	418.90	409.72	416.10	412.10	413.56

- Chan and Grant (2016) showed that cDIC tends to choose over-fitted models while mDIC work better in general
- Model 3 appropriate!

UHASSELT I-BioStat KU LEUVEN
Interuniversity Institute for Biostatistics and statistical Bioinformatics

References

Arellano-Valle, R., Bolfarine, H. and Lachos, V. (2007). Bayesian inference for skew-normal linear mixed models, *Journal of Applied Statistics* **34**(6): 663–682.
URL: *http://dx.doi.org/10.1080/02664760701236905*

Celeux, G., Forbes, F., Robert, C. and Titterington, D. (2006). Deviance information criteria for missing data models, *Bayesian Analysis* **1**: 651–706.

Chan, J. and Grant, A. (2016). On the observed-data deviance information criterion for volatility modeling, *J Financ Economet* **14**(4): 772–802.

Hu, P.and Tsiatis, A. and Davidian, M. (1998). Estimating the parameters in the cox models when covariates variables are measured with error, *Biometrics,* **54**: 14071419.

Lachos, V., Ghosh, P. and Arellano-Valle, R. (2010). Likelihood based inference for skew-normal/independent linear mixed model, *Statistica Sinica* **20**: 303–322.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion), *Journal of the Royal Statistical Society: Series B* **64**(4): 583–616.

Verbeke, G. and Lesaffre, E. (1997). A linear mixed-effect model with heterogeneity in the random-effects population, *Journal of the American Statistical Association* **91**(3): 217–221.

Zhang, D. and Davidian, M. (2001). Liner mixed models with flexible distributions of random effects for longitudinal data, *Biometrics* **57**: 795–802.

Interuniversity Institute for Biostatistics and statistical Bioinformatics

Thank you for your attention.
oludaresamuel.ariyo@kuleuven.be

UHASSELT I-BioStat KU LEUVEN
Interuniversity Institute for Biostatistics and statistical Bioinformatics