THE GENERALIZED LINEAR MIXED MODEL LEADING TERMS

Matt Wand

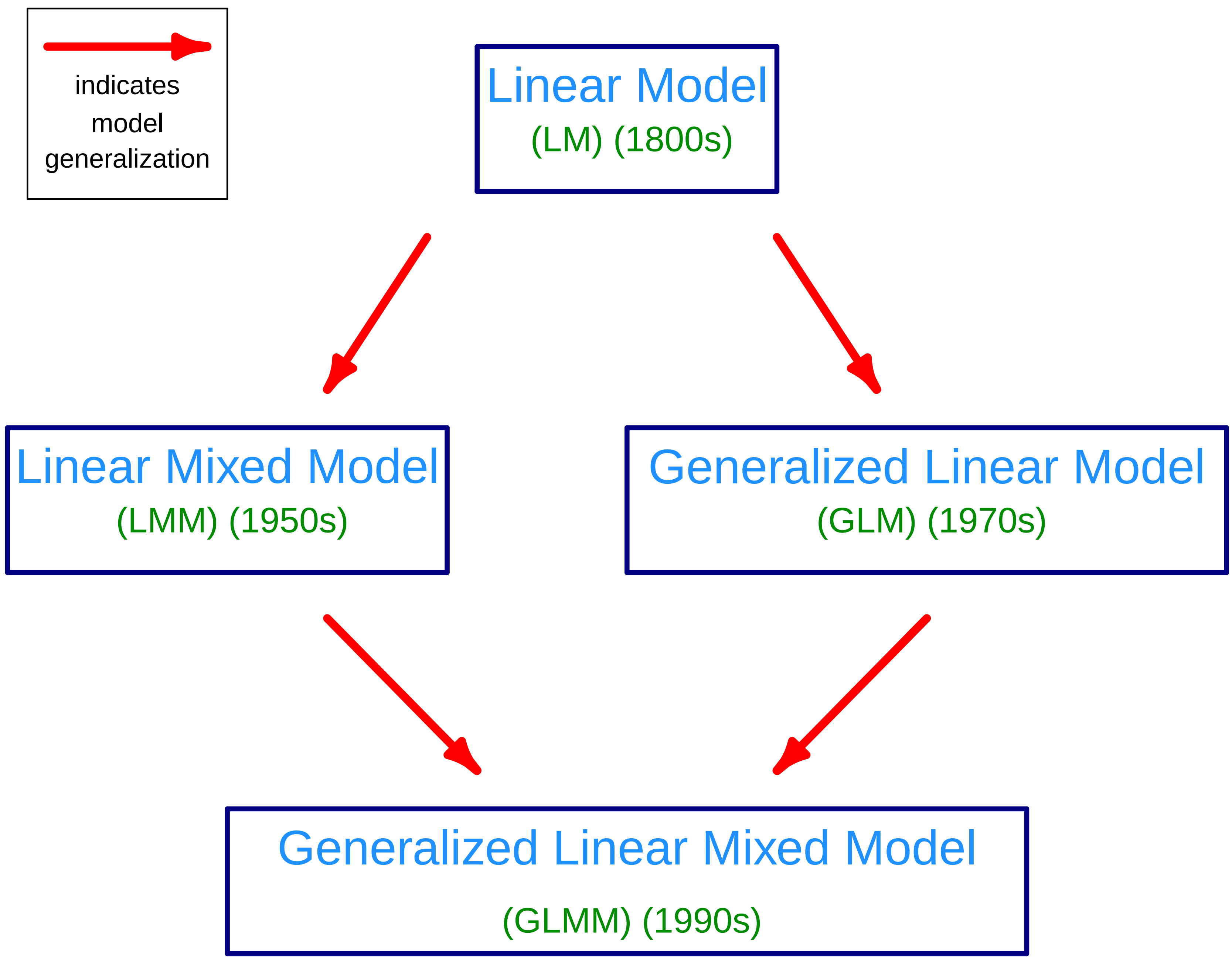
University of Technology Sydney, Australia

joint work with Jiming Jiang, (Univ. California, Davis, U.S.A.)
Aishwarya Bhaskaran (Macquarie University, Australia)
and Luca Maestrini (Australian National University, Australia)

LINEAR MODELS

are a MAINSTAY of

data analysis, statistics,
data science, machine learning.

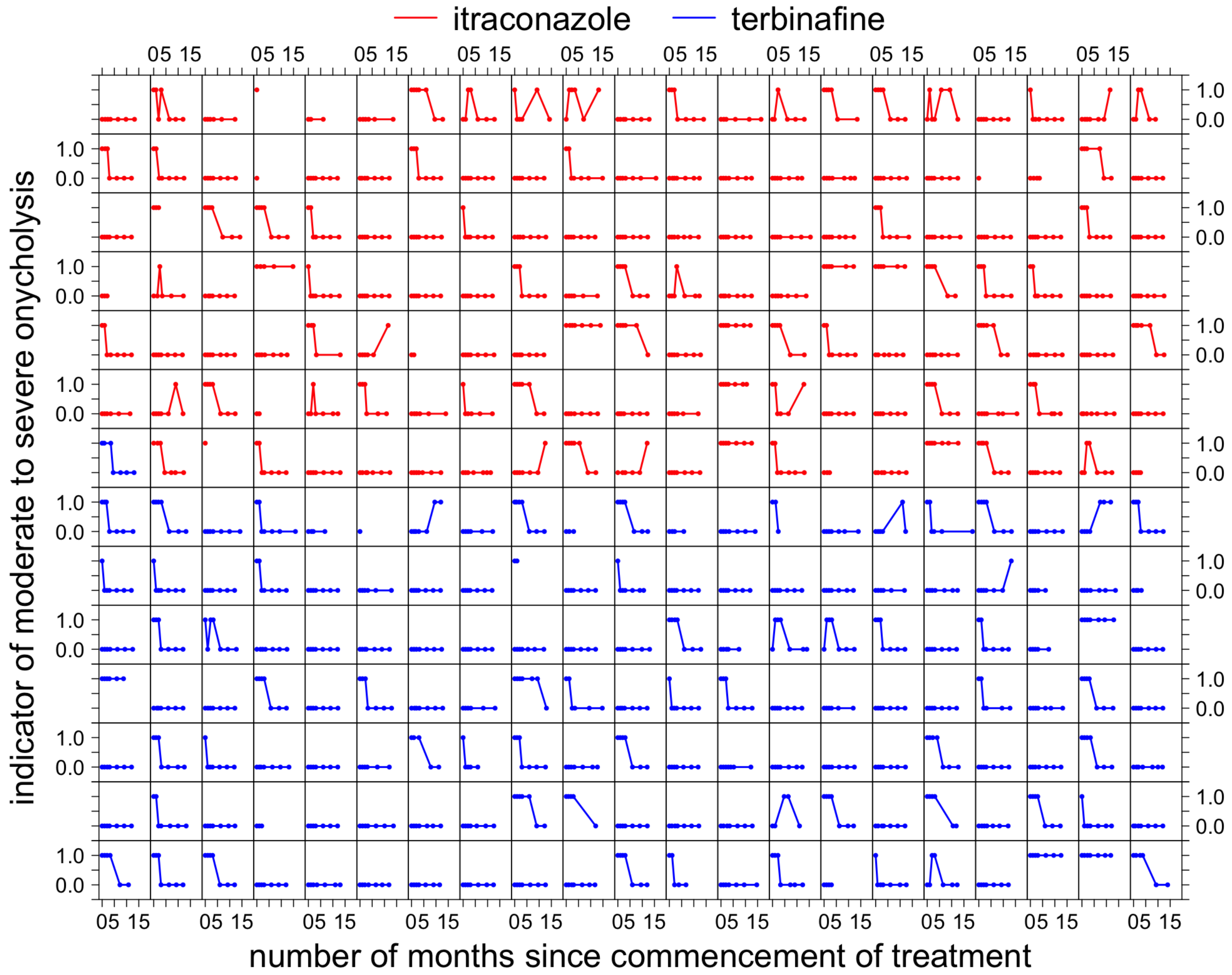

indicates model generalization
Linear Model
(LM) (1800s)
Linear Mixed Model
(LMM) (1950s)
Generalized Linear Model
(GLM) (1970s)
Generalized Linear Mixed Model
(GLMM) (1990s)

Linear Model Effects Structures

model	(conditional) mean response
LM (linear model)	$\boldsymbol{X\beta}$
LMM (linear mixed model)	$\boldsymbol{X\beta + Zu}, \quad \boldsymbol{u} \sim (\mathbf{0}, \boldsymbol{\Sigma})$
GLM (generalized linear model)	$\boldsymbol{g(X\beta)}$
GLMM (generalized linear mixed model)	$\boldsymbol{g(X\beta + Zu)}, \quad \boldsymbol{u} \sim (\mathbf{0}, \boldsymbol{\Sigma})$

COMING UP NEXT...

a data set that benefits from GLMM analysis.



itraconazole
terbinafine
05 15
1.0
0.0
indicator of moderate to severe onycholysis
number of months since commencement of treatment

Generalized Linear Mixed Models as a Topic

Based on a Web of Science analysis for

topic = “generali*ed linear mixed model*”

year	no. papers
1991	1
1993	3
2021	1005

Generalized Linear Mixed Models as a Topic

Based on a Web of Science analysis for

topic = "generali*ed linear mixed model*"

year	no. papers
1991	1
1993	3
2021	1005

$\Longrightarrow$ in the last three decades GLMM has gone from

an emerging area

to

almost 3 new papers EVERY DAY.

Maximum Likelihood Estimation and Asymptotic Variances

TOY EXAMPLE: $\boldsymbol{X_1, \ldots, X_n \overset{\text{ind.}}{\sim} N\Big(\log(\theta^0), 1\Big).}$

$\boldsymbol{\widehat{\theta}} =$ maximum likelihood estimator of $\boldsymbol{\theta^0 = \exp(\overline{X}).}$

EXACT VARIANCE FOR THIS TOY EXAMPLE:

$$\text{Var}(\boldsymbol{\widehat{\theta}}) = \boldsymbol{(\theta^0)^2\{\exp(2/n) - \exp(1/n)\}.}$$

Maximum Likelihood Estimation and Asymptotic Variances

TOY EXAMPLE: $\boldsymbol{X_1, \ldots, X_n \overset{\text{ind.}}{\sim} N\Big(\log(\theta^0), 1\Big).}$

$\widehat{\boldsymbol{\theta}} =$ maximum likelihood estimator of $\boldsymbol{\theta^0 = \exp(\overline{X}).}$

EXACT VARIANCE FOR THIS TOY EXAMPLE:

$$\text{Var}(\widehat{\boldsymbol{\theta}}) = \boldsymbol{(\theta^0)^2\{\exp(2/n) - \exp(1/n)\}.}$$

<u>ASYMPTOTIC</u> VARIANCE FOR THIS TOY EXAMPLE:

$$\text{Asy.Var}(\widehat{\boldsymbol{\theta}}) = \boldsymbol{\frac{(\theta^0)^2}{n}.}$$

Linear Model Asymptotic Variance Example

$$Y_i | X_i \overset{\text{ind.}}{\sim} N\left(\beta_0^0 + \beta_1^0 X_i, (\sigma^2)^0\right), \quad 1 \le i \le n,$$

$X_1, \ldots, X_n$ indep. and identically distributed as X.

IMPORTANT: THE UNIQUE DETERMINISTIC LEADING TERMS BEING SOUGHT!

Linear Model Asymptotic Variance Example

$$Y_i | X_i \overset{\text{ind.}}{\sim} N\left(\beta_0^0 + \beta_1^0 X_i, (\sigma^2)^0\right), \quad 1 \le i \le n,$$

$X_1, \ldots, X_n$ indep. and identically distributed as X.

IMPORTANT: THE <u>UNIQUE DETERMINISTIC</u> LEADING TERMS BEING SOUGHT!

ANSWER FOR β_1: $$\text{Asy.Var}(\widehat{\beta}_1) = \frac{(\sigma^2)^0}{\text{Var}(X)\, n}.$$

EXAMPLE: $X \sim \text{Gamma}(7, 1), \quad (\sigma^2)^0 = 4$

$$\Longrightarrow \text{Asy.Var}(\widehat{\beta}_1) = \frac{(4/7)}{n}$$ PRECISELY

Linear Model Asymptotic Variance Example

$$Y_i | X_i \overset{\text{ind.}}{\sim} N\left(\beta_0^0 + \beta_1^0 X_i, (\sigma^2)^0\right), \quad 1 \le i \le n,$$

$X_1, \ldots, X_n$ indep. and identically distributed as X.

IMPORTANT: THE UNIQUE DETERMINISTIC LEADING TERMS BEING SOUGHT!

ANSWER FOR β_1: $$\text{Asy.Var}(\widehat{\beta}_1) = \frac{(\sigma^2)^0}{\text{Var}(X)\, n}.$$

EXAMPLE: $X \sim \text{Gamma}(7, 1), \quad (\sigma^2)^0 = 4$

$$\Longrightarrow \text{Asy.Var}(\widehat{\beta}_1) = \frac{(4/7)}{n}$$ PRECISELY,

NOT $\frac{(4/7) - 10^{-100}}{n}$ or $\frac{(4/7) + 10^{-100}}{n}$.

Generalized Linear Model Example

$$\boldsymbol{Y_i | X_i} \overset{\text{ind.}}{\sim} \text{Poisson}\Big(\mathbf{exp}\left(\beta_0^0 + \beta_1^0\, \boldsymbol{X_i}\right)\Big), \quad \mathbf{1} \leq \boldsymbol{i} \leq \boldsymbol{n},$$

$$\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{ind.}}{\sim} \text{Uniform}(-\mathbf{1}, \mathbf{1}).$$

$$\Longrightarrow \quad \text{Asy.Var}(\widehat{\beta}_1) = \frac{(\beta_1^0)^3 \,\mathbf{sinh}(\beta_1^0)}{\mathbf{exp}(\beta_0^0)\big\{\,\mathbf{sinh}^2(\beta_1^0) - (\beta_1^0)^2\big\}\, \boldsymbol{n}}.$$

EXAMPLE: If $\beta_0^0 = \mathbf{0.34}$ and $\beta_1^0 = \mathbf{0.82}$ then

$$\Longrightarrow \quad \text{Asy.Var}(\widehat{\beta}_1) \; = \; \frac{\mathbf{2.1778665681\ldots}}{\boldsymbol{n}}.$$

Generalized Linear Mixed Model Example

$$\boldsymbol{Y_{ij}|X_{ij},U_i} \overset{\text{ind.}}{\sim} \text{Poisson}\Big(\mathbf{exp}\left(\boldsymbol{\beta_0^0+U_i+\beta_1^0\,X_{ij}}\right)\Big), \quad \boldsymbol{1 \le i \le m,}$$

$$\boldsymbol{X_{ij}} \overset{\text{ind.}}{\sim} \text{Uniform}(\boldsymbol{-1,1}), \quad \boldsymbol{U_i} \overset{\text{ind.}}{\sim} \boldsymbol{N\big(0,(\sigma^2)^0\big),} \qquad \boldsymbol{1 \le j \le n.}$$

Generalized Linear Mixed Model Example

$$Y_{ij}|X_{ij}, U_i \overset{\text{ind.}}{\sim} \text{Poisson}\Big(\exp\left(\beta_0^0 + U_i + \beta_1^0 X_{ij}\right)\Big), \quad 1 \le i \le m,$$

$$X_{ij} \overset{\text{ind.}}{\sim} \text{Uniform}(-1,1), \quad U_i \overset{\text{ind.}}{\sim} N\left(0, (\sigma^2)^0\right), \qquad 1 \le j \le n.$$

THE SITUATION IN JANUARY 2020...

$$\text{Asy.Var}(\widehat{\beta}_0) = \frac{?}{m}.$$

$$\text{Asy.Var}(\widehat{\beta}_1) = \frac{?}{mn}.$$

THE SITUATION IN JUNE 2020

(THANKS TO THE SPEAKER AND HIS CO-AUTHORS)

$$\text{Asy.Var}(\widehat{\beta}_0) = \frac{(\boldsymbol{\sigma}^2)^0}{\boldsymbol{m}}.$$

$$\text{Asy.Var}(\widehat{\beta}_1) = \frac{(\beta_1^0)^3 \sinh(\beta_1^0)}{\exp(\beta_0^0 + \frac{1}{2}(\boldsymbol{\sigma}^2)^0)\left\{\sinh^2(\beta_1^0) - (\beta_1^0)^2\right\} \boldsymbol{m}\,\boldsymbol{n}}.$$

Example: $\beta_0^0 = -1.46$, $\beta_1^0 = 2.11$, $(\boldsymbol{\sigma}^2)^0 = 6.66$

$$\text{Asy.Var}(\widehat{\beta}_0) = \frac{6.66}{\boldsymbol{m}} \;\; ; \;\; \text{Asy.Var}(\widehat{\beta}_1) = \frac{2.0487282409\ldots}{\boldsymbol{mn}}$$

Random Intercept and Slopes Extension

$$Y_{ij}|X_{ij}, U_{0i}, U_{1i} \overset{\text{ind.}}{\sim} \text{Bernoulli}\left(\text{expit}\left(\beta_0^0 + U_{0i} + (\beta_1^0 + U_{1i})\, X_{1ij} + \beta_2^0\, X_{2ij}\right)\right),$$

$$\text{expit}(x) \equiv \frac{e^x}{1 + e^x}, \quad \begin{bmatrix} U_{0i} \\ U_{1i} \end{bmatrix} \overset{\text{ind.}}{\sim} N(0, \Sigma^0).$$

$$\text{Asy.Var}(\widehat{\beta}_0) = \frac{(\Sigma)_{11}^0}{m},$$

$$\text{Asy.Var}(\widehat{\beta}_1) = \frac{(\Sigma)_{22}^0}{m}, \quad \text{(note how simple!)}$$

$$\text{Asy.Var}(\widehat{\beta}_2) = \frac{C_2}{mn} \quad (C_2 \text{ next page}).$$

$$\frac{1}{C_2} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{(4\pi)^{-1}|\Sigma^0|^{-1/2} \exp\left\{ -\frac{1}{2}[u_0\ u_1](\Sigma^0)^{-1}[u_0\ u_1]^T \right\} du_0 du_1}{\left(\left[E\left\{ \dfrac{\begin{bmatrix} 1 & X_1 & X_2 \\ X_1 & X_1^2 & X_1 X_2 \\ X_2 & X_1 X_2 & X_2^2 \end{bmatrix}}{\cosh\left(\beta_0^0 + u_0 + (\beta_1^0 + u_1) X_1 + \beta_2^0 X_2\right) + 1} \right\} \right]^{-1} \right)_{33}}$$

$$\frac{1}{C_2} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{(4\pi)^{-1}|\Sigma^0|^{-1/2}\exp\left\{-\frac{1}{2}[u_0\ u_1](\Sigma^0)^{-1}[u_0\ u_1]^T\right\}du_0du_1}{\left(\left[E\left\{\dfrac{\begin{bmatrix} 1 & X_1 & X_2 \\ X_1 & X_1^2 & X_1X_2 \\ X_2 & X_1X_2 & X_2^2 \end{bmatrix}}{\cosh\left(\beta_0^0 + u_0 + (\beta_1^0 + u_1)X_1 + \beta_2^0 X_2\right) + 1}\right\}\right]^{-1}\right)_{33}}$$

- GOOD NEWS: Studentization is possible via method of moments estimation.
- NOT SO GREAT NEWS: We're stuck with a bivariate integral.

The Simpler Random Intercepts-Only Model

$$Y_{ij}|X_{ij}, U_{0i}, U_{1i} \overset{\text{ind.}}{\sim} \text{Bernoulli}\left(\text{expit}\left(\beta_0^0 + U_{0i} + \beta_1^0 X_{1ij} + \beta_2^0 X_{2ij}\right)\right),$$

$$U_{0i} \overset{\text{ind.}}{\sim} N\left(0, (\sigma^2)^0\right)$$

$$\begin{aligned} \text{Asy.Var}(\widehat{\beta}_0) &= \frac{(\Sigma)_{11}^0}{m}, \\ \text{Asy.Var}(\widehat{\beta}_1) &= \frac{C_1}{mn}, \\ \text{Asy.Var}(\widehat{\beta}_2) &= \frac{C_2}{mn}. \end{aligned}$$

Now C_1 and C_2 only require univariate quadrature.

(Potential) Benefits of GLMM Leading Term Theory

- Confidence intervals
- Wald hypothesis tests
- Optimal design
- Sample size calculations

Excerpt from Diggle *et. al*

Analysis of Longitudinal Data

Let z_p denote the pth quantile of a standard Gaussian distribution and $d = \beta_{1B} - \beta_{1A}$ be the meaningful difference of interest. With n fixed and known, the number of subjects per group that are needed to achieve type I error rate α and power P, is

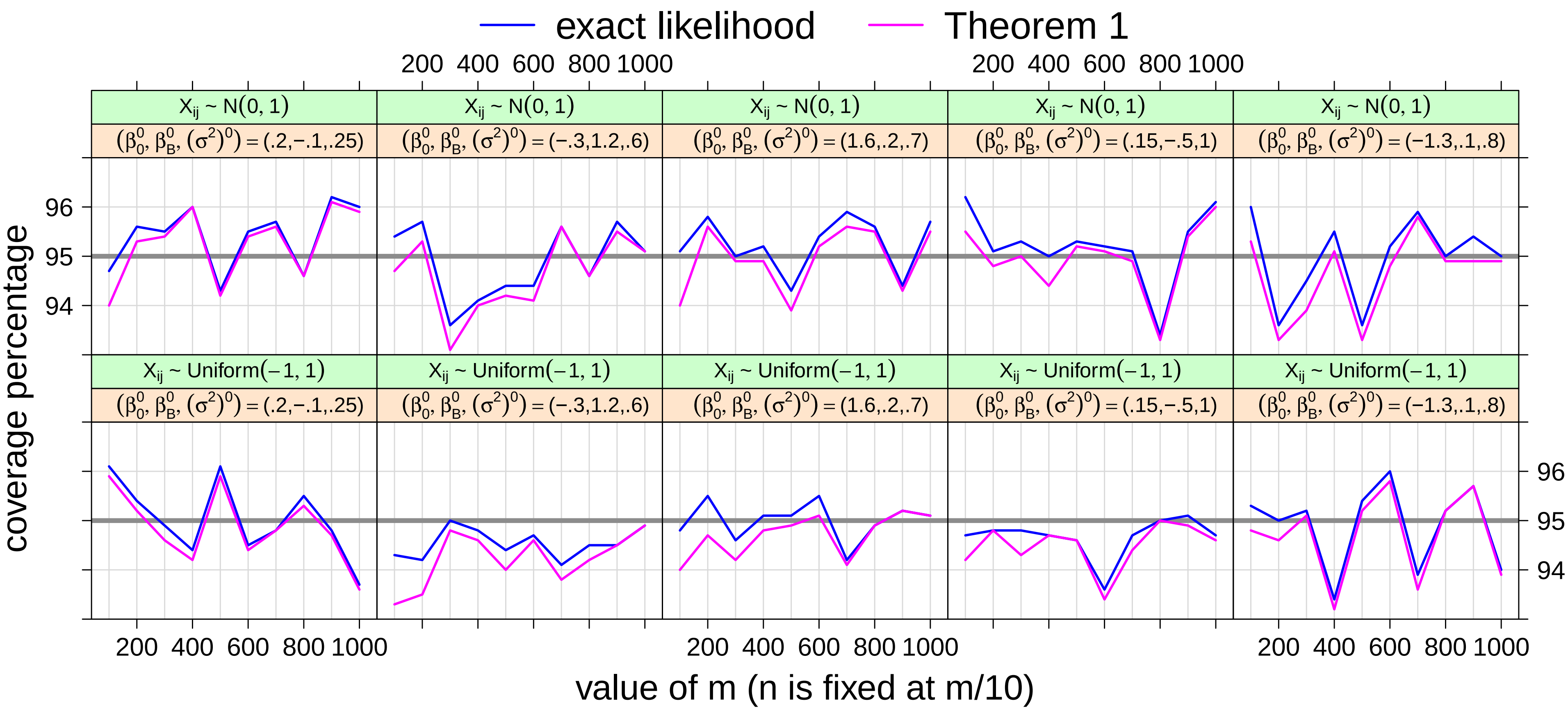
$$m = \frac{2(z_\alpha + z_Q)^2 \sigma^2 (1-\rho)}{n s_x^2 d^2}, \tag{2.4.1}$$

where $Q = 1 - P$ and $s_x^2 = \sum_j (x_j - \bar{x})^2/n$, the within-subject variance of the x_j.

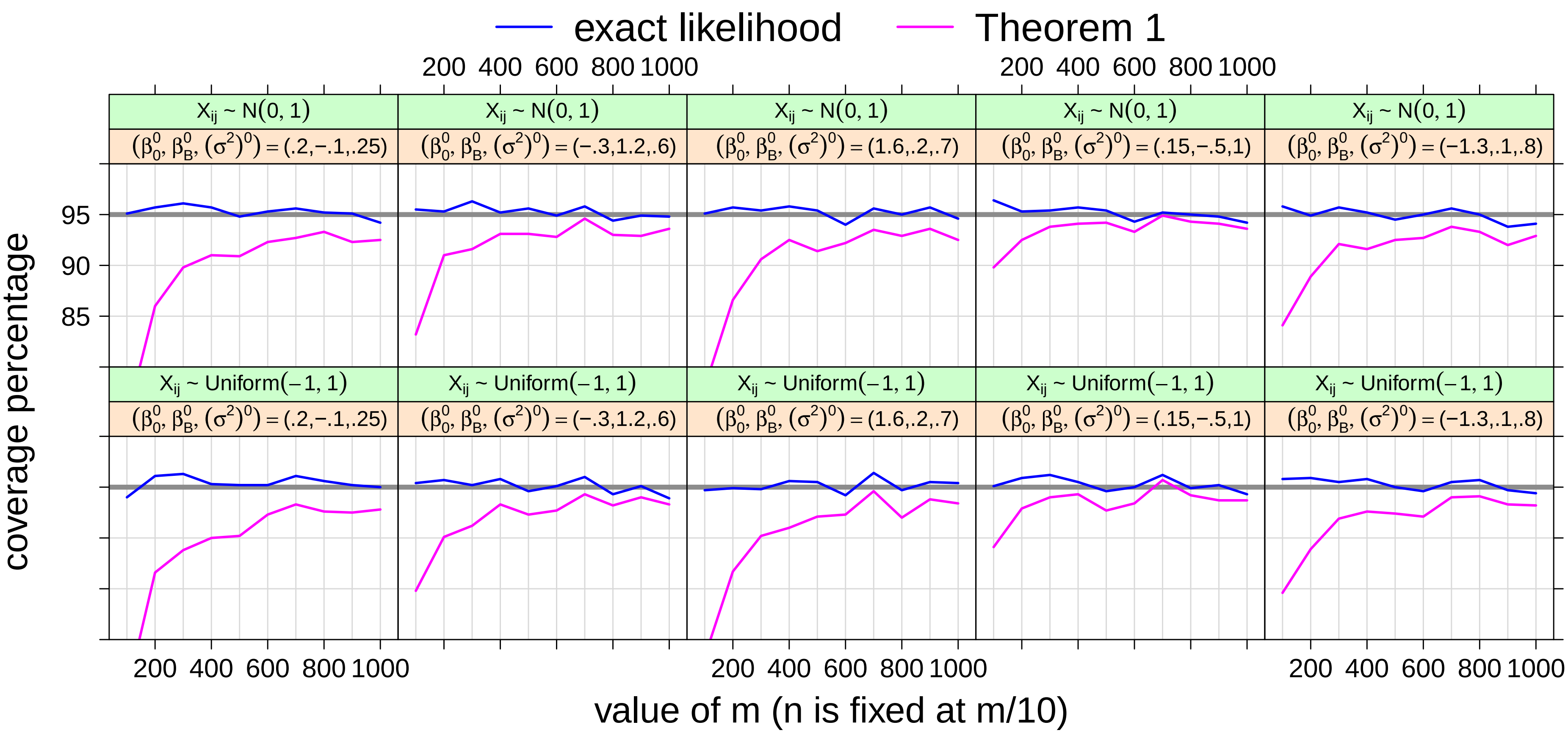
To illustrate, consider a hypothetical clinical trial on the effect of a new treatment in reducing blood pressure. Three visits, including the baseline visit, are planned at years 0, 2, and 5. Thus, $n = 3$ and $s_x^2 = 4.22$. For type I error rate $\alpha = 0.05$, power $P = 0.8$ and smallest meaningful difference $d = 0.5\,\text{mmHg/year}$, the table below gives the number of subjects needed for both treated and control groups for some selected values of ρ and σ^2.

	σ^2		
ρ	100	200	300
0.2	313	625	937
0.5	195	391	586
0.8	79	157	235

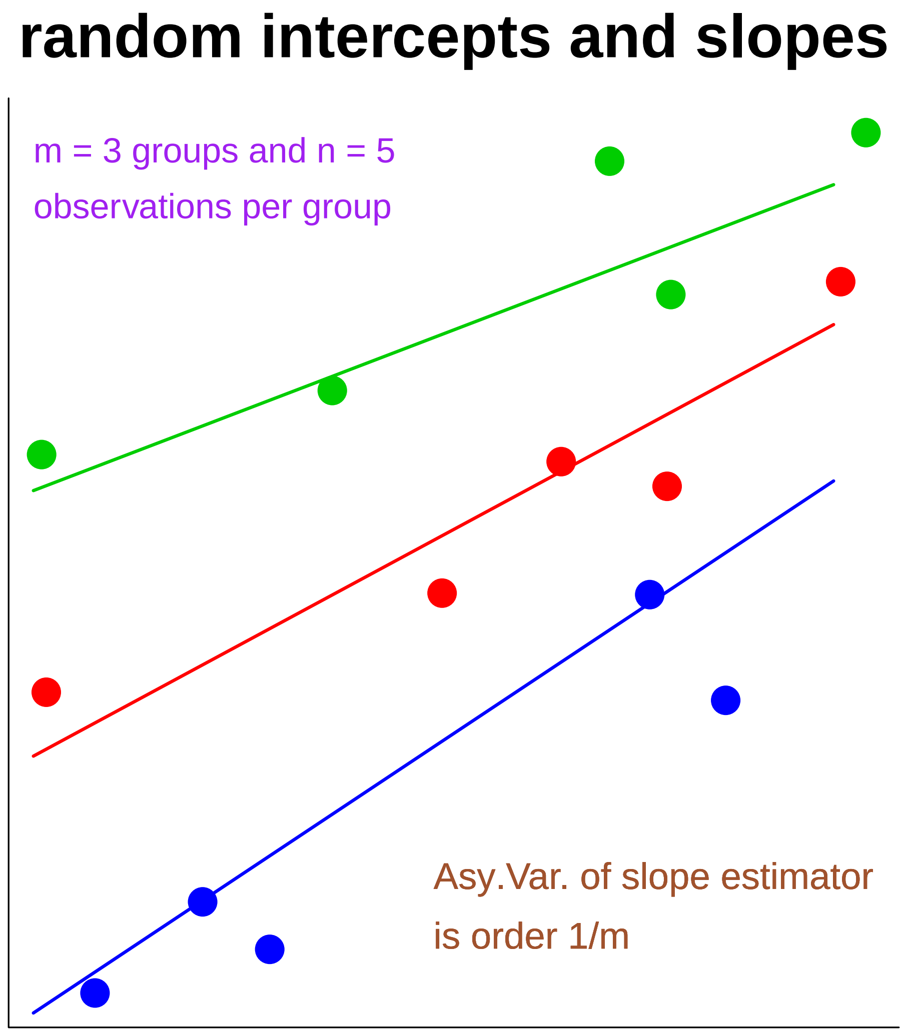
Empirical Coverage Results for an Order $(mn)^{-1}$ Parameter



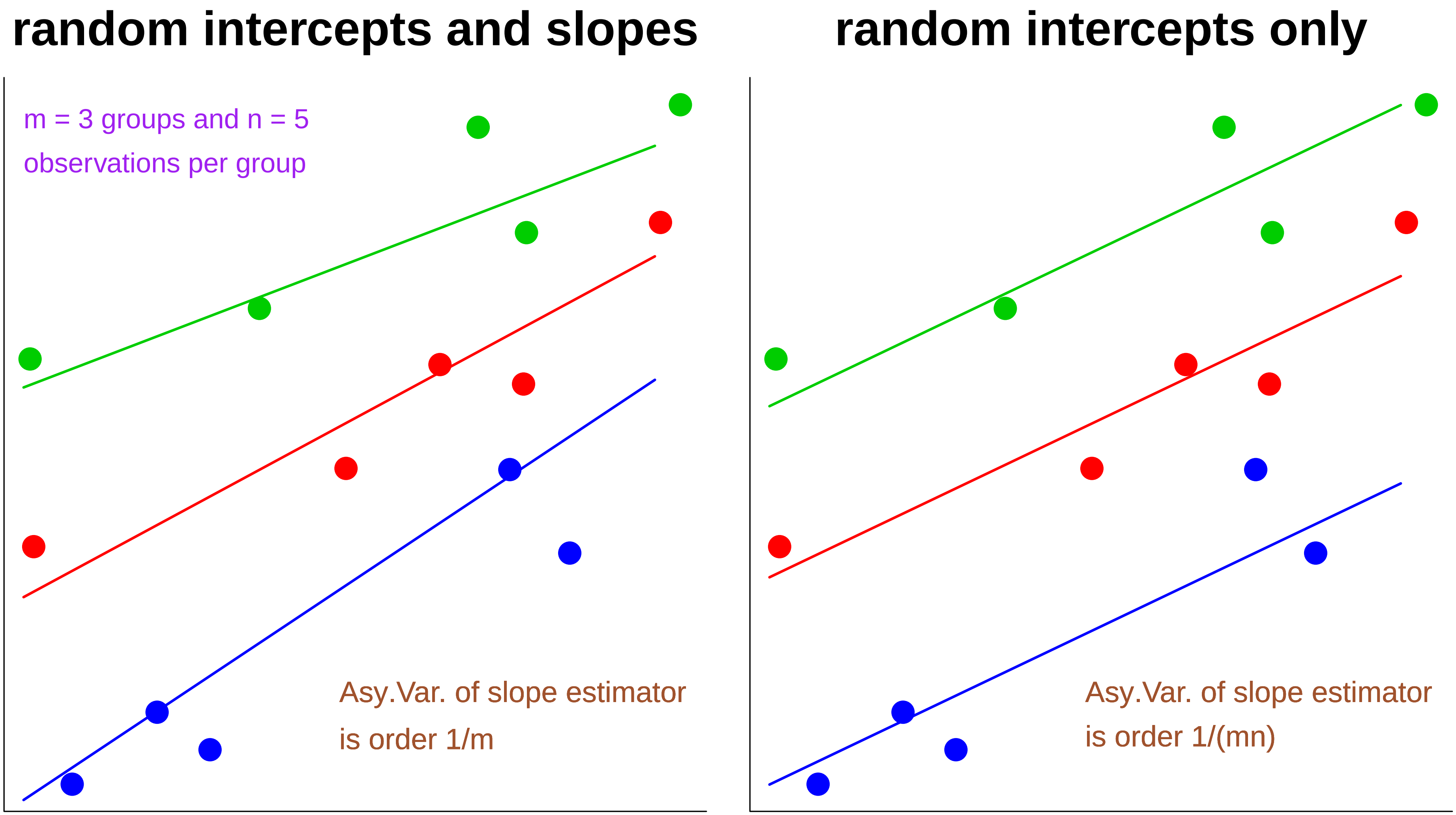
Empirical Coverage Results for an Order m^{-1} Parameter



Some Rates of Convergence Intuition



Some Rates of Convergence Intuition



Our Most General GLMM Leading Term Result

The linear predictor is $(\boldsymbol{\beta}^0_{\mathsf{A}} + \boldsymbol{U}_i)^T \boldsymbol{X}_{\mathsf{A}ij} + (\boldsymbol{\beta}^0_{\mathsf{B}})^T \boldsymbol{X}_{\mathsf{B}ij}$.

The random effects vector is $\boldsymbol{U}_i \overset{\text{ind.}}{\sim} N\big(\boldsymbol{0}, \boldsymbol{\Sigma}^0\big)$.

The dispersion parameter is ϕ^0; appears in the log-likelihood via $d(\phi)$.

The matrix $\boldsymbol{C}_{\mathsf{B}}$ has a complicated expression (á la $\boldsymbol{C}_2$ a few slides back).

The matrix $\boldsymbol{D}^+_{d_{\mathsf{A}}}$ has a universal form arising from the vech operator.

Some moment conditions are imposed.

$$\sqrt{m}\begin{bmatrix} \widehat{\boldsymbol{\beta}}_{\mathsf{A}} - \boldsymbol{\beta}^0_{\mathsf{A}} \\ \sqrt{n}\Big(\widehat{\boldsymbol{\beta}}_{\mathsf{B}} - \boldsymbol{\beta}^0_{\mathsf{B}}\Big) \\ \text{vech}(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}^0) \\ \sqrt{n}(\widehat{\phi} - \phi^0) \end{bmatrix} \xrightarrow{\mathcal{D}} N\left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ 0 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}^0 & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \phi^0 \boldsymbol{C}_{\mathsf{B}} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & 2\boldsymbol{D}^+_{d_{\mathsf{A}}}(\boldsymbol{\Sigma}^0 \otimes \boldsymbol{\Sigma}^0)\boldsymbol{D}^{+T}_{d_{\mathsf{A}}} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \dfrac{1}{2d'(\phi^0)/\phi^0 + d''(\phi^0)} \end{bmatrix}\right).$$

Gaussian Variational Approximation Connection

In Hall, Pham, Wand & Wang (Ann. Statist., 2011) we treated Poisson mixed models with log-likelihood:

$$\ell(\boldsymbol{\beta},\sigma^2)=\sum_{i=1}^{m}\log\int_{-\infty}^{\infty}\exp\left\{\sum_{j=1}^{n}(Y_{ij}u-e^{\beta_0+\beta_1 X_{ij}+u})-\frac{u^2}{2\sigma^2}\right\}du+\text{CLOSED FORM}$$

The ith of these integrals can be written as

$$\int_{-\infty}^{\infty}\exp\left\{\sum_{j=1}^{n}(Y_{ij}u-e^{\beta_0+\beta_1 X_{ij}+u})-\frac{u^2}{2\sigma^2}\right\}\frac{\frac{1}{\sqrt{2\pi\lambda_i}}e^{-\frac{1}{2}(u-\mu_i)^2/\lambda_i}}{\frac{1}{\sqrt{2\pi\lambda_i}}e^{-\frac{1}{2}(u-\mu_i)^2/\lambda_i}}\,du$$

$$=\sqrt{2\pi\lambda_i}\,E_{\tilde{U}_i}\left[\exp\left\{\sum_{j=1}^{n}\left(Y_{ij}\tilde{U}_i-e^{\beta_0+\beta_1 X_{ij}+\tilde{U}_i}\right)-\frac{\tilde{U}_i^2}{2\sigma^2}+\frac{(\tilde{U}_i-\mu_i)^2}{2\lambda_i}\right\}\right]$$

where $E_{\tilde{U}_i}$ denotes expectation with respect to the random variable $\tilde{U}_i\sim N(\mu_i,\lambda_i)$. LOOKS LIKE A JOB FOR JENSEN'S INEQUALITY...

Gaussian Variational Approximation Connection (continued)

Application of Jensen's inequality leads to

$$\ell(\boldsymbol{\beta},\boldsymbol{\sigma}^2) \;\geq\; \underline{\ell}(\boldsymbol{\beta},\boldsymbol{\sigma}^2,\boldsymbol{\mu},\boldsymbol{\lambda})$$ WHICH HAS A CLOSED FORM!

$$(\underline{\widehat{\boldsymbol{\beta}}},\underline{\widehat{\boldsymbol{\sigma}}}^2) = (\boldsymbol{\beta},\boldsymbol{\sigma}^2) \text{ component of } \underset{\boldsymbol{\beta},\boldsymbol{\sigma}^2,\boldsymbol{\mu},\boldsymbol{\lambda}}{\operatorname{argmax}}\,\underline{\ell}(\boldsymbol{\beta},\boldsymbol{\sigma}^2,\boldsymbol{\mu},\boldsymbol{\lambda}).$$

Back in 2011 we proved that

$$\text{Asy.Var}(\underline{\widehat{\beta}_0}) = \frac{(\boldsymbol{\sigma}^2)_0}{m}, \quad \text{Asy.Var}(\underline{\widehat{\beta}_1}) = \frac{C_1}{mn}, \quad \text{Asy.Var}(\underline{\widehat{\boldsymbol{\sigma}}}^2) = \frac{2\{(\boldsymbol{\sigma}^2)_0\}^2}{m}.$$

WHICH ARE THE SAME ASYMPTOTIC VARIANCES FOR EXACT MAXIMUM LIKELIHOOD! AND (ELEVEN YEARS LATER) GAUSSIAN VARIATIONAL APPROXIMATION IS FULLY EFFICIENT!

References For This Talk So Far

Jiang, J., Wand, M.P. and Bhaskaran, A. (2022).

Usable and precise asymptotics for generalized linear mixed model analysis and design.

Journal of the Royal Statistical Society, Series B.

Bhaskaran, A. and Wand, M.P. (2023).

Dispersion parameter extension of precise generalized linear mixed model asymptotics.

Statistics and Probability Letters.

Two-Term Asymptotic Variances (joint with Luca Maestrini)

$$\boldsymbol{\beta} \equiv \left[\begin{array}{c} \boldsymbol{\beta}_{\mathsf{A}} \\ \boldsymbol{\beta}_{\mathsf{B}} \end{array} \right], \ \boldsymbol{X}_{ij} \equiv \left[\begin{array}{c} \boldsymbol{X}_{\mathsf{A}ij} \\ \boldsymbol{X}_{\mathsf{B}ij} \end{array} \right] \implies \left(\boldsymbol{\beta} + \left[\begin{array}{c} \boldsymbol{U}_i \\ \boldsymbol{0} \end{array} \right] \right)^T \boldsymbol{X}_{ij} \text{ is linear predictor}$$

GAUSSIAN RESPONSE CASE (LMM):

$$\mathrm{Cov}\big(\widehat{\boldsymbol{\beta}}|\boldsymbol{\mathcal{X}}\big) = \frac{\mathbf{1}}{\boldsymbol{m}} \left[\begin{array}{cc} \boldsymbol{\Sigma}^0 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{array} \right] + \frac{\phi\big\{\boldsymbol{E}(\boldsymbol{X}\boldsymbol{X}^T)\big\}^{-1}\{\mathbf{1} + \boldsymbol{o}_{\boldsymbol{P}}(\mathbf{1})\}}{\boldsymbol{mn}}$$

Two-Term Asymptotic Variances (joint with Luca Maestrini)

$$\boldsymbol{\beta} \equiv \begin{bmatrix} \boldsymbol{\beta}_{\mathsf{A}} \\ \boldsymbol{\beta}_{\mathsf{B}} \end{bmatrix}, \ \boldsymbol{X}_{ij} \equiv \begin{bmatrix} \boldsymbol{X}_{\mathsf{A}ij} \\ \boldsymbol{X}_{\mathsf{B}ij} \end{bmatrix} \implies \left(\boldsymbol{\beta} + \begin{bmatrix} \boldsymbol{U}_i \\ \mathbf{0} \end{bmatrix}\right)^T \boldsymbol{X}_{ij} \text{ is linear predictor}$$

GAUSSIAN RESPONSE CASE (LMM):

$$\text{Cov}(\widehat{\boldsymbol{\beta}}|\boldsymbol{\mathcal{X}}) = \frac{1}{m}\begin{bmatrix} \boldsymbol{\Sigma}^0 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix} + \frac{\phi\{\boldsymbol{E}(\boldsymbol{X}\boldsymbol{X}^T)\}^{-1}\{1 + o_P(1)\}}{mn}$$

GENERAL RESPONSE CASE (GLMM):

$$\text{Cov}(\widehat{\boldsymbol{\beta}}|\boldsymbol{\mathcal{X}}) = \frac{1}{m}\begin{bmatrix} \boldsymbol{\Sigma}^0 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix} + \frac{\phi\boldsymbol{K}\{1 + o_P(1)\}}{mn}.$$

We found the expression for $\boldsymbol{K}$, but it ain't pretty.

LATE BREAKING NEWS

Our two-term paper accepted at

BIOMETRIKA

during the week before last!!!

The Clincher

Part of a table from the *Biometrika*-accepted paper concerning

sample size calculations for 90% power
with $\boldsymbol{\alpha = 0.05}$ in a logistic mixed model set-up

for particular values of model parameters and $\boldsymbol{n = 20}$:

	1-term var.	2-term var.
Minimum m:	54	304
Power estimate:	32.8	88.5
Power conf. int.:	$(29.9, 35.7)$	$(86.5, 90.5)$

An Additional 2024 Reference For This Talk

Maestrini, L., Bhaskaran, A. and Wand, M.P. (2024).

Second term improvement to generalised linear mixed model asymptotics.

Biometrika.

CROSSED

RATHER THAN

NESTED

RANDOM EFFECTS

This is CURRENT joint research with Swarnadip Ghosh, Art Owen (Stanford University, U.S.A.) and Jiming Jiang (University of California, Davis, U.S.A.).

What Makes Crossed Harder Than Nested?

As per an e-mail message from Art Owen to speaker (April 2022):

- how recent even a consistency for GLMM proof was (Jiang, Ann. Statist., 2013),
- the superlinear costs,
- the non-existence of a bootstrap.

MORE ON THE CROSSED VS. NESTED DIFFICULTIES IN NEXT FEW SLIDES...

Crossed Linear Mixed Model Example

$$\boldsymbol{Y_{ii'j}|X_{ii'j},U_i,U'_{i'}} \stackrel{\text{ind.}}{\sim} \boldsymbol{N}\Big(\beta_0^0+\boldsymbol{U_i}+\boldsymbol{U'_{i'}}+\beta_1^0\,\boldsymbol{X_{ii'j}},\boldsymbol{\sigma_\varepsilon^2}\Big), 1 \le i \le m, 1 \le i' \le m'$$

$$\boldsymbol{X_{ii'j}} \stackrel{\text{ind.}}{\sim} \boldsymbol{X}, \quad \boldsymbol{U_i} \stackrel{\text{ind.}}{\sim} \boldsymbol{N}\big(\boldsymbol{0},(\boldsymbol{\sigma^2})^0\big), \quad \boldsymbol{U'_{i'}} \stackrel{\text{ind.}}{\sim} \boldsymbol{N}\big(\boldsymbol{0},\{(\boldsymbol{\sigma'})^2\}^0\big), \quad \boldsymbol{1 \le j \le n.}$$

THE SITUATION IN MARCH 2022...

$$\text{Asy.Var}(\widehat{\beta}_0) = ?$$

$$\text{Asy.Var}(\widehat{\beta}_1) = ?$$

EVEN IN THE GAUSSIAN RESPONSE CASE!

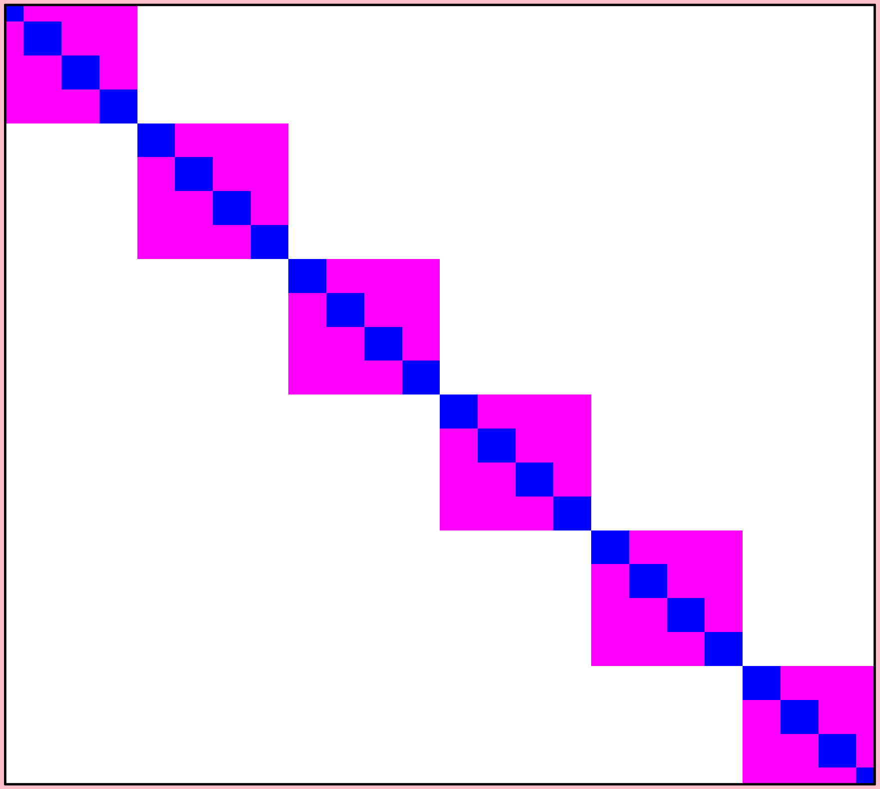
Linear Mixed Models in Their Most General Form

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\varepsilon}$$

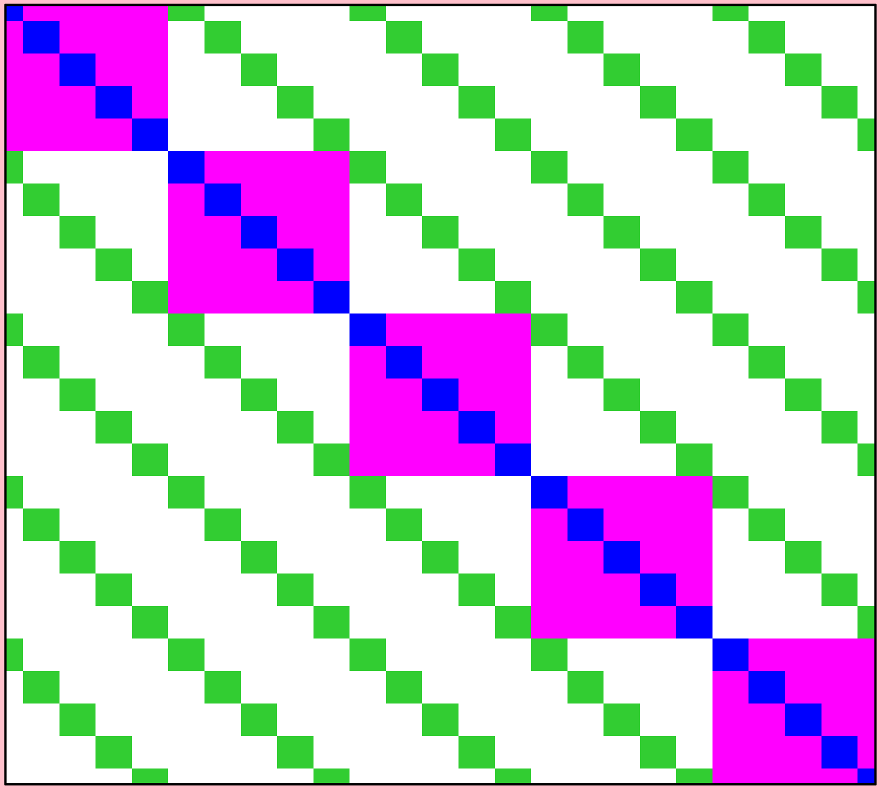
$$\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} \sim \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{G} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{R} \end{bmatrix} \right)$$

$$\mathrm{Cov}(\widehat{\boldsymbol{\beta}}) = \left(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1} \text{ where } \boldsymbol{V} \equiv \boldsymbol{ZGZ}^T + \boldsymbol{R}.$$

The V matrix for a nested case



The V matrix for a crossed case



Crossed Linear Mixed Model Example

$$Y_{ii'j}|X_{ii'j}, U_i, U'_{i'} \overset{\text{ind.}}{\sim} N\left(\beta_0^0 + U_i + U'_{i'} + \beta_1^0\, X_{ii'j}, \sigma_\varepsilon^2\right), 1 \le i \le m, 1 \le i' \le m'$$

$$X_{ii'j} \overset{\text{ind.}}{\sim} X, \quad U_i \overset{\text{ind.}}{\sim} N(0, (\sigma^2)^0), \quad U'_{i'} \overset{\text{ind.}}{\sim} N(0, \{(\sigma')^2\}^0), \quad 1 \le j \le n.$$

THE SITUATION IN MAY 2022...

$$\text{Asy.Var}(\widehat{\beta}_0) = \frac{\sigma^2}{m} + \frac{(\sigma')^2}{m'}.$$

$$\text{Asy.Var}(\widehat{\beta}_1) = \frac{\sigma_\varepsilon^2}{mm'n\text{Var}(X)}.$$

A Very General and NEW Crossed Random Effects Leading Terms "Result"

$$\boldsymbol{Y}_{ii'}|\boldsymbol{U}_i,\boldsymbol{U}'_{i'},\boldsymbol{X}_{\mathrm{A}ii'},\boldsymbol{X}_{\mathrm{B}ii'} \overset{\text{ind.}}{\sim} N\Big((\boldsymbol{\beta}^0_{\mathrm{A}}+\boldsymbol{U}_i+\boldsymbol{U}'_{i'})^T\boldsymbol{X}_{\mathrm{A}ii'}+(\boldsymbol{\beta}^0_{\mathrm{B}})^T\boldsymbol{X}_{\mathrm{B}ii'},(\boldsymbol{\sigma}^2)^0\boldsymbol{I}\Big)$$

$$\boldsymbol{U}_i \overset{\text{ind.}}{\sim} N\big(\boldsymbol{0},\boldsymbol{\Sigma}^0\big),\quad 1\le i\le m,\quad \boldsymbol{U}'_{i'} \overset{\text{ind.}}{\sim} N\big(\boldsymbol{0},(\boldsymbol{\Sigma}')^0\big),\quad 1\le i'\le m'.$$

Under some assumptions including $m = O(m')$ and $m' = O(m)$:

$$\left[\begin{array}{c}
\left\{\dfrac{\boldsymbol{\Sigma}^0}{m}+\dfrac{(\boldsymbol{\Sigma}')^0}{m'}\right\}^{-1/2}\Big(\widehat{\boldsymbol{\beta}}_{\mathrm{A}}-\boldsymbol{\beta}^0_{\mathrm{A}}\Big)\\
\left\{\dfrac{(\boldsymbol{\sigma}^2)^0\boldsymbol{C}_B}{mm'n}\right\}^{-1/2}\Big(\widehat{\boldsymbol{\beta}}_{\mathrm{B}}-\boldsymbol{\beta}^0_{\mathrm{B}}\Big)\\
\left\{\dfrac{2\boldsymbol{D}^+_{d_{\mathrm{A}}}(\boldsymbol{\Sigma}^0\otimes\boldsymbol{\Sigma}^0)\boldsymbol{D}^{+T}_{d_{\mathrm{A}}}}{m}\right\}^{-1/2}\mathrm{vech}\big(\widehat{\boldsymbol{\Sigma}}-\boldsymbol{\Sigma}^0\big)\\
\left\{\dfrac{2\boldsymbol{D}^+_{d_{\mathrm{A}}}((\boldsymbol{\Sigma}')^0\otimes(\boldsymbol{\Sigma}')^0)\boldsymbol{D}^{+T}_{d_{\mathrm{A}}}}{m'}\right\}^{-1/2}\mathrm{vech}\big(\widehat{\boldsymbol{\Sigma}}'-(\boldsymbol{\Sigma}')^0\big)\\
\left\{\dfrac{2((\boldsymbol{\sigma}^2)^0)^2}{mm'n}\right\}^{-1/2}\big(\widehat{\boldsymbol{\sigma}}^2-(\boldsymbol{\sigma}^2)^0\big)
\end{array}\right] \xrightarrow{\mathcal{D}} N(\boldsymbol{0},\boldsymbol{I}).$$

THE CROSSED CASE REQUIRES
MUCH MORE WORK
THAN THE NESTED CASE

AND THIS IS STILL JUST
FOR THE GAUSSIAN
RESPONSE (LMM) CASE!

Closing Summary

- Leading term asymptotics has several practical benefits including confidence intervals, optimal design and sample size calculations.
- There is also the attraction of simple-to-digest summaries of the behaviour of the model parameter estimators.
- Despite 30 years of GLMM research and widespread use, leading term asymptotics has been a

MAJOR GAP.

- We show how to BRIDGE THIS MAJOR GAP.

FOR RELEVANT PAPERS AND MORE CHECK OUT...

matt-wand.utsacademics.info