欠測のあるデータにおけるpopulation-averaged及びsubject-specificアプローチの性能評価

多田　圭佑
サノフィ株式会社

研究開発部門　医薬開発本部

統計解析・プログラミング部　統計解析室

土居　正明　駒嵜　弘

Performance evaluation of population-averaged and subject-specific approach with missing data

Keisuke Tada
Sanofi Biostatistics Biostatistics & Programming
Clinical Sciences & Operations Research & Development

要旨:

はじめにpopulation-averaged及びsubject-specific modelについてや2種類の手法の関係を説明し、例として欠測のある経時的な計数データに対して2種類の手法の性能を評価する。

キーワード:

population-averaged model, subject-specific model, binary data, count data, genmod procedure, gee procedure, glimmix procedure

本日の発表内容

- ✓2種類のモデル
  - Subject-Specific (SS) Model
  - Population-Averaged (PA) Model
- ✓ SSとPAとの関係
  - SSからPAへの変換方法
    - 恒等リンク
    - ログリンク
    - ロジットリンク
- ✓ シミュレーションによる性能評価
  - SSモデルによるデータセットの作成方法
  - 結果と考察
- ✓ まとめ

Subject-Specific (SS) Model

✓ モデル式[1][2][6]

$$g(E[y_{ij}|\boldsymbol{\gamma}_i]) = \mathbf{x}_{ij}^t\boldsymbol{\beta} + \mathbf{z}_{ij}^t\boldsymbol{\gamma}_i$$

$$\boldsymbol{\gamma}_i \sim N(\mathbf{0}, \mathbf{D})$$

e.g. 線形混合効果モデル(LMM)
一般化線形混合効果モデル(GLMM)

g：リンク関数
β：固定効果
γ：変量効果
i：個人
j：時点

✓ 特徴

1. 個人ごとの変量効果で条件付けされた期待値をモデル化
2. βは個人ごとの変量効果で条件付けられた薬効
3. 臨床試験で有用

Population-Average (PA) Model

✓ モデル式

$$g^*(E[y_{ij}]) = \mathbf{x}_{ij}^t \boldsymbol{\beta}^*$$

注意：
先ほどのβとは
異なります！

e.g. 線形モデル(LM)
一般化推定方程式による推定法(GEE法)

✓ 特徴

1. 応答の周辺期待値をモデル化
2. β*は集団の平均的薬効
3. 条件付分布を仮定する必要なし
4. 疫学に有用

SSとPAとの関係

SS : GLMM
PA : GEE法

- ✓ SSからPAへの変換は比較的容易[2][3]
  ⇒PAとしての薬剤効果を求めるためにSSモデルでの推定結果をPAとして読み替えることが可能
  - ➢ 条件付期待値の期待値を取ることで、応答の周辺期待値を求める。

$$E[y_{ij}] = E[E[(y_{ij}|\boldsymbol{\gamma}_i)]]$$

$$= \int g^{-1}(\mathbf{x}_{ij}^t\boldsymbol{\beta} + \mathbf{z}_{ij}^t\boldsymbol{\gamma}_i)dF(\boldsymbol{\gamma}_i)$$

  - ➢ リンク関数に具体的な関数を当てはめ、積分計算を行う。

SSからPAへの変換

SS : GLMM
PA : GEE法

✓ 恒等リンクの場合

$g(\mu) = \mu \Rightarrow g(\mu)^{-1} = \mu$ のとき

$$E[y_{ij}] = \int (\mathbf{x}_{ij}^t \boldsymbol{\beta} + \mathbf{z}_{ij}^t \boldsymbol{\gamma}_i) dF(\boldsymbol{\gamma}_i)$$

$$= \mathbf{x}_{ij}^t \boldsymbol{\beta}$$

SS : $E[y_{ij}] = \mathbf{x}_{ij}^t \boldsymbol{\beta}$

PA : $E[y_{ij}] = \mathbf{x}_{ij}^t \boldsymbol{\beta}^*$

$\boxed{\beta^* = \beta}$

PAで推定する固定効果はSSと同様のスケールで解釈可能！

SSからPAへの変換

SS : GLMM
PA : GEE法

✓ ログリンクの場合

$$g(\mu) = \log(\mu) \Rightarrow g(\mu)^{-1} = \exp(\mu)$$ のとき

$$E[y_{ij}] = \int \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta} + \mathbf{z}_{ij}^t \boldsymbol{\gamma}_i) dF(\boldsymbol{\gamma}_i)$$

$$= \exp\left(\mathbf{x}_{ij}^t \boldsymbol{\beta} + \mathbf{z}_{ij}^t D \mathbf{z}_{ij}/2\right)$$

SS : $\log(E[y_{ij}]) = \mathbf{x}_{ij}^t \boldsymbol{\beta} + \mathbf{z}_{ij}^t D \mathbf{z}_{ij}/2$

PA : $\log(E[y_{ij}]) = \mathbf{x}_{ij}^t \boldsymbol{\beta}^*$

$\boxed{\beta^* = \beta}$ ※切片は異なるので注意

PAで推定する固定効果はSSと同様のスケールで解釈可能！

SSからPAへの変換

SS : GLMM
PA : GEE法

✓ ロジットリンクの場合

$$g(\mu) = \mathrm{logit}(\mu) \Rightarrow g(\mu)^{-1} = \frac{1}{1+\exp(-\mu)}$$

のとき

SS : $\mathrm{logit}(E[y_{ij}]) \approx a(\mathbf{D})\mathbf{x}_{ij}^t\boldsymbol{\beta}$

$$a(\mathbf{D}) = |c^2\mathbf{D}\mathbf{z}_{ij}\mathbf{z}_{ij}^t + \mathbf{I}|^{-q/2} \qquad \frac{1}{1+\exp(-t)} \approx \Phi\left(\frac{16\sqrt{3}}{15\pi}\cdot t\right) \qquad c = \frac{16\sqrt{3}}{15\pi}$$

PA : $\mathrm{logit}(E[y_{ij}]) = \mathbf{x}_{ij}^t\boldsymbol{\beta}^*$

$$\boxed{\beta^* = a(D)\beta}$$

PAで推定する固定効果はSSの定数倍の変換が必要！

ここまでのまとめ

SS : GLMM
PA : GEE法

Subject-Specific (SS) Model

Population-Averaged (PA) Model

- ✓ 評価の対象が異なる。
- ✓ SSからPAへ変換して、PAとして解釈可能
  - ➢ 恒等リンク、ログリンクの場合：
    固定効果は同様のスケールで解釈可能
  - ➢ ロジットリンクの場合：
    固定効果は定数倍（変量効果の分散に依存）に変換が必要

シミュレーションによる性能評価

SS : GLMM
PA : GEE法

- ✓ 共通設定
  - 応答:計数データ
  - 試験デザイン:
    プラセボ対照並行群間比較試験(検証試験)
  - 時点:5つ(ベースライン含む)
  - 欠測のメカニズム:MAR
  - 欠測のパターン:monotone
- ✓ 条件
  - 例数 : 200 / 群
  - 欠測の量:多い(最終時点約50%脱落)、
    少ない(最終時点約15%脱落)
  - 時点間相関:0.2, 0.6, 0.8

欠測の量を少なく、時点間の相関を
約0.6とした場合の結果を主に示します。

シミュレーションによる性能評価

SS : GLMM
PA : GEE法

✓ 比較対象

➢ **最終時点の薬剤効果を表す固定効果（推定値とその相対バイアス）**

今回は1000回です。

$$r.bias = \begin{cases} \frac{1}{n}\sum_{i=1}^{n} \frac{\hat{\beta}_i - \beta_t}{\beta_t} * 100 & (\beta_t \neq 0) \\ \frac{1}{n}\sum_{i=1}^{n} \hat{\beta}_i & (\beta_t = 0) \end{cases}$$

n ：シミュレーション回数　$\hat{\beta}_i$ ：i番目の薬剤効果の推定値　β_t ：真の薬剤効果

➢ **αエラー（有意水準:0.05）**

◆ 全ての時点における真の薬剤効果（群間差）を「なし」としたときの有意となった試験の割合

➢ **検出力**

◆ 真の薬剤効果（群間差）を「あり」としたときの有意となった試験の割合

シミュレーションによる性能評価

SS : GLMM
PA : GEE法

✓ データセット作成方法（SSベース）

$$
\begin{aligned}
\log(E[Y_j|\gamma]) = \beta_0 &+ \beta_1 * x_1 \\
&+ (\beta_{21} + \beta_{31} * x_3) * I(j=2) \\
&+ (\beta_{22} + \beta_{32} * x_3) * I(j=3) \\
&+ (\beta_{23} + \boxed{\beta_{33}} * x_3) * I(j=4) \\
&+ \gamma
\end{aligned}
$$

評価の対象

$$H_0 : \beta_{33} = 0$$

$x_1 : baseline \sim Po(4)$ ベースライン

$$x_3 = \begin{cases} 0 & (placebo) \\ 1 & (active) \end{cases}$$ 群（治験薬服用の有無）

$\gamma \sim N(0, \sigma_\gamma^2)$ 被験者ごとの変量効果

パラメータ	値
β_0	0.2
β_1	0.15
β_{21}	-0.05
β_{22}	-0.1
β_{23}	-0.2
β_{31}	-0.3
β_{32}	-0.4
β_{33}	-0.5
σ^2	0.75

シミュレーションによる性能評価

SS : GLMM
PA : GEE法

- ✓ 欠測の発生方法
  - ➢ 観測確率モデル
    共変量として、1時点前の応答と同様の値と
    各時点の効果を入れる。

$$\mathrm{logit}(\lambda_j) = \alpha_0 + \alpha_1 * Y_{j-1} + \alpha_{21} * I(j=3) + \alpha_{22} * I(j=4)$$

パラメータ	値
α_0	3
α_1	-0.1
α_{21}	0.3
α_{22}	0.1

  - ➢ 時点j-1において観測され、
    時点jにおいても観測する確率

$$Pr\{R_j = 1 | R_{j-1} = 1, Y, \boldsymbol{\alpha}\} = \lambda_j$$

要約統計量（真の薬剤効果あり）

Group	TP	mean	std	median	obs rate
Placebo	1	4.0071	1.9968	3.9740	1.0000
	2	3.0960	3.4703	2.0445	1.0000
	3	2.8754	3.1295	2.0125	0.9498
	4	2.6733	2.8816	1.9800	0.8940
	5	2.3723	2.5786	1.7985	0.8368
Active	1	4.0040	1.9982	3.9770	1.0000
	2	3.0843	3.4069	2.0470	1.0000
	3	2.1296	2.4188	1.3780	0.9493
	4	1.8063	2.1051	1.0260	0.8975
	5	1.4627	1.7722	1.0000	0.8461

ベースライン

時点3から
欠測を発生

※小数第5位以下は切り捨てて表示しています。

シミュレーションによる性能評価

SS : GLMM
PA : GEE法

- ✓ 比較① （PA内での比較）
  - ➢ 欠測ありなしのデータに対して
    - ◆ GEE法
  - ➢ 欠測のあるデータに対して
    - ◆ GEE法
    - ◆ wGEE法[4][5]

```
proc gee data = XXXX plot = (all);

    class arm(ref="0") caseid tp;
    missmodel prevar tp / type = obslevel;
    model var_m = bl arm tp arm * tp / dist=poisson link = log;
    repeated subject=caseid / corr=cs corrw maxit = 100;

run;
```

比較① 結果

✓ 最終時点の薬剤効果の推定値とバイアス

真の薬剤効果	手法	beta33	n	mean	se	r. bias
なし	欠測のないデータにGEE	0	1000	0.0029	0.1092	0.0029
	欠測のあるデータにGEE	0	1000	0.0029	0.1102	0.0029
	wGEE	0	1000	0.0018	0.1256	0.0018
あり	欠測のないデータにGEE	-0.5	1000	-0.5022	0.1147	0.4477
	欠測のあるデータにGEE	-0.5	1000	-0.4918	0.1174	-1.6348
	wGEE	-0.5	1000	-0.4978	0.1307	-0.4443

比較① 結果

✓ αエラーと検出力

手法	αエラー	検出力
欠測のないデータにGEE	0.055	0.995
欠測のあるデータにGEE	0.049	0.982
wGEE	0.055	0.948

- αエラー
  本条件下ではおおよそ5%に近い。

- 検出力
  wGEEがやや低下か（SEが大きいためか）。
  ※シミュレーションを増やして検討が必要である。

比較① 結果

✓ 観測確率モデル (wGEE)

		真の薬剤効果なし			真の薬剤効果あり		
parameter	true	n	mean	r. bias	n	mean	r. bias
alpha0	3	1000	3.0335	1.1172	1000	3.0172	0.5743
alpha1	-0.1	1000	-0.1003	0.2932	1000	-0.0978	-2.1619
alpha21	0.3	1000	0.3020	0.6720	1000	0.3165	5.4939
alpha22	0.1	1000	0.0799	-20.1159	1000	0.1086	8.5607

- 観測確率モデルの推定にバイアスが入る。
  ⇒応答変数モデルの推定にバイアスが入る。
  ⇒αエラー・検出力に影響する可能性あり。

- 真の薬剤効果なしのalpha22に大きめのバイアスあり。

比較① 条件ごとの比較

✓ 時点間相関弱い、やや強い、強い条件で比較(SE)

真の薬剤効果	手法	相関0.2	相関0.6	相関0.8
なし	欠測のないデータにGEE	0.0728	0.1092	0.1427
	欠測のあるデータにGEE	0.0790	0.1102	0.1319
	wGEE	0.0791	0.1256	0.1673
あり	欠測のないデータにGEE	0.0834	0.1147	0.1472
	欠測のあるデータにGEE	0.0904	0.1174	0.1391
	wGEE	0.0905	0.1307	0.1708

- 相関が大きくなるに従い、wGEE法はGEE法よりSEが大きくなる傾向がある。

比較① 条件ごとの比較

✓ 時点間相関弱い、やや強い、強い条件で比較(αエラーと検出力)

	αエラー			検出力		
手法	相関0.2	相関0.6	相関0.8	相関0.2	相関0.6	相関0.8
欠測のないデータにGEE	0.053	0.055	0.063	1.000	0.995	0.899
欠測のあるデータにGEE	0.052	0.049	0.070	0.999	0.982	0.909
wGEE	0.053	0.055	0.096	0.999	0.948	0.796

- 相関が強い場合、全体的にαエラーがインフレするか。
- 検出力等を考慮した、現実的な例数による追加シミュレーションが必要か。

比較① 条件ごとの比較

✓ 欠測多め（最終時点の脱落約50%）、相関0.6の結果

真の薬剤効果	手法	beta33	n	mean	se	r. bias
なし	欠測のないデータにGEE	0	1000	0.0029	0.1092	0.0029
	欠測のあるデータにGEE	0	1000	0.0024	0.1137	0.0024
	wGEE	0	967	-0.0038	0.2901	-0.0038
あり	欠測のないデータにGEE	-0.5	1000	-0.5022	0.1147	0.4477
	欠測のあるデータにGEE	-0.5	1000	-0.4196	0.1213	-16.0706
	wGEE	-0.5	972	-0.4493	0.2165	-10.1335

比較① 考察

- ✓ 時点間相関が強い場合、wGEEではαエラーのインフレが起きそう(外れ値が増えるためか、しかしシミュレーション回数が少ないため結論づけるのは軽率)。
- ✓ 本条件下では欠測が多いとwGEEのバイアスも大きくなる。
- ✓ 欠測が多少あってもGEEはそこまで性能は落ちないか。
- ✓ wGEEにおいて、重みが大きい場合は薬剤効果の推定値は極端な値になる可能性がある。
  ⇒現在のprocedureでは算出された重みはデータセットには出力できず、plot=histgramで視覚的にのみ確認可能。

※注意事項
エラーが出力されたものは結果に含めていない。
収束しなくてもnoteのみ表示されて推定値は出力されるため、
収束しなかった推定値は手作業で除外している。

シミュレーションによる性能評価

SS : GLMM
PA : GEE法

- ✓ 比較② SS内での比較
  - ➢ **欠測ありなしのデータに対してGLMM**
    - ◆ LAPLACE法
    - ◆ RSPL法
    - ◆ QUAD法

```
proc glimmix data = XXXX method = laplace;

    class arm(ref="0") caseid tp;
    model var = bl arm tp arm * tp / s dist=poisson link = log;
    random int / subject = caseid type=un;

run;
```

比較② 結果

✓ GLMMの3種類の手法の比較（欠測なし）

			欠測なし			
真の薬剤効果	手法	beta33	n	mean	se	r. bias
なし	laplace	0	1000	0.0029	0.1001	0.0029
	quad	0	1000	0.0029	0.1004	0.0029
	rspl	0	981	0.0036	0.0982	0.0036
あり	laplace	-0.5	1000	-0.4998	0.1067	-0.0417
	quad	-0.5	1000	-0.4997	0.1070	-0.0568
	rspl	-0.5	979	-0.4892	0.1047	-2.1618

□ RSPLのバイアスが比較的大きく、収束しない場合がある。

比較② 結果

✓ GLMMの3種類の手法の比較（欠測あり）

			欠測あり			
真の薬剤効果	手法	beta33	n	mean	se	r. bias
なし	laplace	0	1000	0.0030	0.1066	0.0030
	quad	0	1000	0.0031	0.1069	0.0031
	rspl	0	990	0.0026	0.1047	0.0026
あり	laplace	-0.5	1000	-0.5000	0.1139	0.0059
	quad	-0.5	1000	-0.5000	0.1142	-0.0057
	rspl	-0.5	993	-0.4894	0.1119	-2.1256

□ 欠測の有無に影響されにくいか。

比較② 結果

✓ αエラーと検出力

手法	欠測	αエラー	検出力
laplace	なし	0.052	0.998
	あり	0.046	0.988
quad	なし	0.050	0.998
	あり	0.045	0.988
rspl	なし	0.051	0.997
	あり	0.042	0.988

- 本発表の条件下では大きな違いはないといえる。
- 強いて挙げるとすると、RSPLがたまに収束せず、バイアスも比較的高めである。

比較② 条件を変更

✓ 時点間相関を弱くした(約0.2)結果

			欠測あり			
真の薬剤効果	手法	beta33	n	mean	se	r. bias
なし	laplace	0	1000	0.0028	0.0795	0.0028
	quad	0	1000	0.0028	0.0794	0.0028
	rspl	0	465	-0.0005	0.0789	-0.0005
あり	laplace	-0.5	1000	-0.5003	0.0910	0.0542
	quad	-0.5	1000	-0.5003	0.0909	0.0523
	rspl	-0.5	296	-0.5029	0.0906	0.5773

- LAPLACE及びQUADの性能は同等か。
- RSPLの収束する回数が極端に下がり、バイアスも増えている。

比較② 考察

- ✓ 時点間相関がやや強い場合は、3手法に大きな相違はなさそう。
- ✓ 時点間相関が弱い場合は、method = RSPLの収束が悪く、バイアスも多い。
- ✓ 欠測の量にはあまり依存しないか。

シミュレーションによる性能評価

SS : GLMM
PA : GEE法

- ✓ 比較③ （PAとしての比較）
  - ➢ 欠測のあるデータに対して
    - ◆ GEE法
    - ◆ wGEE法
    - ◆ GLMMの結果をPAへ変換
      - LAPLACE
      - QUAD
      - RSPL

比較③ 結果(r. bias)

※真の薬剤効果ありの場合

- ✓ 時点間相関やや強い、欠測の量少なめ
  GLMM(RSPL) > GEE > wGEE
  > GLMM(LAPLACE, QUAD)

- ✓ 時点間相関やや強い、欠測の量多め
  GEE > wGEE > GLMM(RSPL)
  > GLMM(LAPLACE, QUAD)

- ✓ 時点間相関弱い、欠測の量少なめ
  GLMM(RSPL) > wGEE = GEE
  = GLMM(LAPLACE, QUAD)

※相対バイアスの絶対値を比較

比較③ 結果(αエラー)

- ✓ 時点間相関やや強い、欠測の量少なめ
  wGEE > 0.05 > GEE > GLMM(LAPLACE, QUAD) > GLMM(RSPL)

- ✓ 時点間相関やや強い、欠測の量多め
  wGEE > GEE > 0.05 ≒ GLMM

- ✓ 時点間相関弱い、欠測の量少なめ
  GEE = wGEE = GLMM(LAPLACE, QUAD) = GLMM(RSPL) ≒ 0.05

比較③ 結果(検出力)

- ✓ 時点間相関やや強い、欠測の量少なめ
  GLMM > GEE > wGEE
- ✓ 時点間相関やや強い、欠測の量多め
  GEE > GLMM >> wGEE
- ✓ 時点間相関弱い、欠測の量少なめ
  wGEE = GEE = GLMM

比較③ 考察とまとめ

- ✓ 下記の2種類を比較
  - ➢ PAとして解析したGEE法の結果
  - ➢ SSのモデル・推定方法(GLMM)を用いた上でPAへ変換した結果
- ✓ 正規分布のときのMMRM > wGEEと同様[7][8]に、GLMMのlaplaceやquadが本発表の条件下では、欠測の量や相関の強弱に左右されなく、使用しやすいか。
- ✓ PAを求めたいがなんらかの制限によりPAモデルが使用できないとき、SSモデルで推定・変換してPAとして解釈が可能か。

全体のまとめ

✓2種類のモデリング
- Subject-Specific (SS) model
- Population-Averaged (PA) model
  モデルの立て方、解釈、変換

✓欠測のある経時的な計数データに対する様々な手法のPAにおける比較
- GEE法
- wGEE法
- 一般化線形混合効果モデル(GLMM)の結果をPAに変換

参考文献

[1] Generalized, Linear, and Mixed Models (Wiley Series in Probability and Statistics Charles E. McCulloch et. Al.
[2] Zeger, Scott L., Kung-Yee Liang, and Paul S. Albert. "Models for longitudinal data: a generalized estimating equation approach." *Biometrics* (1988): 1049-1060.
[3] Hu, Frank B., et al. "Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes." *American Journal of Epidemiology* 147.7 (1998): 694-703.
[4] Preisser, John S., Kurt K. Lohman, and Paul J. Rathouz. "Performance of weighted estimating equations for longitudinal binary data with drop - outs missing at random." *Statistics in medicine* 21.20 (2002): 3035-3054.
[5] Guixian Lin, Robert N. Rodriguez SAS Institute Inc. “Weighted Methods for Analyzing Missing Data with the GEE Procedure” Paper SAS166-2015
[6] Lee, Youngjo, and John A. Nelder. "Conditional and marginal models: another view." *Statistical Science* 19.2 (2004): 219-238.
[7] 土居正明, 大浦智紀, 大江基貴, 駒嵜弘, 髙橋文博, 縄田成毅, 藤原正和, 横溝孝明, 横山雄一. (2014). 欠測のあるデータに対する総合的な感度分析と主解析の選択. SASユーザー総会論文集.
[8] 横山雄一 , 横溝孝明, 大浦智紀, 大江基貴. (2015).【日本製薬工業協会シンポジウム】臨床試験の欠測データの取り扱いに関する最近の展開と今後の課題について ー統計手法・estimandと架空の事例に対する流れの整理ー (7)架空の事例2（主解析の選択・例数設計・データの発生方法）