MMRM における重みの取り扱いの検討
— SAS PROC MIXED / GLIMMIX の重み指定の違いと実務上の注意点 —

○江口 凌平[1]、中西 豊支[1]

([1]シミック株式会社 データサイエンス事業本部)

An Examination of Weight Handling in Weighted Mixed Models
— Specification of Weights in SAS PROC MIXED / GLIMMIX and Practical Considerations —

○Ryohei Eguchi[1], Yushi Nakanishi[1]

([1]Data Science Div., CMIC Co., Ltd.)

要旨

反復測定混合モデル（MMRM）は、臨床試験の経時データ解析において標準的な手法として広く用いられている。近年、欠測や選択に起因する偏りへの対応として、観測データの代表性を考慮した重みを組み合わせる重み付き MMRM (IPW-MMRM 等)が注目されている。しかし、SAS の PROC MIXED で指定する WEIGHT ステートメントと、尤度に直接重みを掛ける「尤度重み付け」には数理的・意味合い的に重大な相違がある。本論文では、SAS における重み付けの挙動についてシミュレーションを用いた検証を行い、その違いについて考察する。PROC MIXED の WEIGHT は、観測値または被験者の尤度寄与を直接重み付けするものではなく、残差側（R-side）の分散共分散構造を重み（精度重み（Precision Weight））に応じて変換する。一方、PROC GLIMMIX の OBSWEIGHT オプション等は、個々の観測値の代表性を調整する「尤度重み（Likelihood Weight）」として機能する。さらに、PROC GLIMMIX において観測値固有の重み（尤度重み）を適用する場合には、繰り返し測定の残差分散共分散構造（UN 等）を R-side でモデル化できないという実装上の技術的制約が存在することを解説する。本論文は、実務において重み付き混合モデルを適切に適用するための実践的な指針を提示する。

キーワード：反復測定混合モデル（MMRM）、精度重み付け、尤度重み付け、PROC MIXED、PROC GLIMMIX

1. MMRM 解析において仮定される分散共分散構造

反復測定データ解析（MMRM; Mixed Model for Repeated Measures）では、同一被験者から得られた複数時点の測定値の間に相関があることを前提とする。標準的な MMRM においては、同じ測定スケジュールを持つすべての被験者について、共分散行列は等しいと仮定される。例えば、時点 1、時点 2、時点 3 の 3 時点の評価スケジュールにおいて、被験者 A と被験者 B のそれぞれの残差共分散行列 R_A および R_B は共通であり、非構造化（UN; Unstructured）共分散構造を仮定した場合は以下のように定式化される。

$$R_A = R_B = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}$$

ここで、σ_t^2 は各時点 t における分散、σ_{ts} は時点 t と s の間の共分散を表す。しかし、臨床試験の実データにおいては、被験者ごとに測定の精度やばらつきが異なる場合が想定される。例えば、被験者 B の測定データの分散が被験者 A の 2 倍大きいと仮定する。このとき、測定精度を考慮して解析を補正するために、被験者 A には $W_A = 2$、被験者 B には $W_B = 1$ という重みを与えて調整を行うことができる。これが「精度重み（Precision Weight）」の基本的な考え方であり、測定精度の高い被験者のデータを重視して分散を補正するアプローチとなる。

2. 精度重みづけと尤度重みづけの考え方の違い

2.1 精度重みづけ（Precision Weighting）の数理と特徴

PROC MIXED において WEIGHT ステートメントを指定した場合、指定された重みは、観測値の尤度寄与を直接重み付けするものではなく、残差側（R-side）の分散共分散構造を変換するために用いられる。以下では、この PROC MIXED における WEIGHT ステートメントの数理的な作用を説明する。

観測値数を n、観測値 i (i = 1, …, n)に指定した重みを w_i とする。このとき、重み行列 W は、各観測値の重みを対角成分に配置し、それ以外の成分を 0 とした次の対角行列である。

$$W = \begin{pmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n \end{pmatrix}$$

また、W の逆平方根行列を L とすると、L は次のように表される。

$$L = \begin{pmatrix} \frac{1}{\sqrt{w_1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{w_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{w_n}} \end{pmatrix}$$

すなわち、L の各対角成分は対応する重み w_i の逆平方根となる。通常の線形混合モデルでは、観測値ベクトル Y の分散共分散行列 V は、以下のように表される。

$$V = ZGZ' + R$$

ここで、G はランダム効果 (G-side)の分散共分散行列、R は残差側 (R-side)の分散共分散行列である。REPEATED ステートメントを指定した場合、PROC MIXED の WEIGHT ステートメントは、残差分散共分散行列 R を LRL と変換する。したがって、WEIGHT ステートメントを指定した場合の分散共分散行列 V は、以下のようになる。

$$V = ZGZ' + LRL$$

例えば、R の i 行 i 列の対角成分 R_{ii} は、観測値 i の残差分散を表す。重みを適用した後の観測値 i の残差分散は、以下のようになる。

$$Var(\varepsilon_i, \varepsilon_j) = R_{ij} \,/\, w_i$$

したがって、重み w_i が大きい観測値ほど残差分散は小さくなる。すなわち、重みが大きい観測値は、より測定精度が高い観測値として扱われる。反対に、重みが小さい観測値では残差分散が大きくなり、相対的に測定精度が低い観測値として扱われる。つまり、PROC MIXED の WEIGHT ステートメントは、残差分散共分散行列を重みに応じて変換する精度重み付けとして実装されている。

2.2 尤度重みづけ（Likelihood Weighting）の数理と特徴

疫学や因果推論の解析では、欠測、脱落、治療選択、あるいは標本抽出などにより生じ得る対象集団との構成の違いを考慮するため、各被験者または各観測値に重みを設定することがある。このような重みは、対象集団における個々のデータの頻度、重要度、または代表性を調整することを目的とする。被験者別の尤度重みを考慮した経時データ解析について、PROC GLIMMIX を用いた実装が考えられる。PROC GLIMMIX では、データが階層構造になっている際のモデル(Multilevel model)を実装することが可能である。例えば、地域、施設、世帯、個人など、複数の階層において抽出確率が異なる場合があり、各階層の単位が対象集団において代表する程度を推定に反映するために、階層ごとの重みを用いることがある。一方、通常の臨床試験における経時データ解析では、複雑な標本抽出設計を用いる調査のように、施設の抽出確率や上位階層の抽出確率を反映した重みを用いることは一般的ではない。例えば、施設を第 3 レベル、被験者を第 2 レベル、被験者内の反復測定値を第 1 レベルとみなすことができるが、通常、施設が母集団から抽出される確率に基づいて施設別の重みを設定することはない。臨床試験において重みを用いる場合には、被験者の観測されやすさ、解析対象集団に対する代表性、あるいは欠測・脱落等に関連する要因を考慮し、被験者単位の重みを設定する場面が中心となる。以下では、各被験者に一つの重みを設定し、その重みを当該被験者の全測定時点に共通して適用する場合を考える。

PROC GLIMMIX において尤度重みを適用する場合、目的関数である対数尤度の各被験者または観測値の寄与に対して、直接重み w を乗じる数理的アプローチをとる。被験者 i の反復測定値を y_i とし、被験者 i のデータが対象集団を代表する程度に関係する確率をπ_iとする。このとき、当該単位に与える重み w_i は、観測確率の逆数として次のように定義される。

$$w_i = 1/\pi_i$$

ここで、観測確率π_iが小さい、すなわち対象集団に存在するにもかかわらずデータとして観測されにくい単位ほど、大きな重みが与えられる。この重みは、その単位が対象集団において何単位分を代表するかを表す量として解釈できる。全体の対数擬似尤度は、各単位の対数尤度への寄与を重み付きで加算したものとして、次のように表される。

$$\log\big(p(y)\big) = \sum w_i \log(p(y_i))$$

ここでの$p(y_i)$は i 番目のデータが得られる確率(尤度)、$\log(p(y_i))$は対数尤度への寄与を表す。この式は、各単位の対数尤度に対して重みを直接乗じることを意味するため、重みが大きい単位は、パラメータ推定においてより大きな寄与を持つ。例えば、観測確率が 0.25 であれば、その重みは 4 となる。この単位は、対象集

団において概念的には 4 単位分を代表するものとして、対数尤度への寄与が 4 倍される。このように、重みが大きい単位ほど推定への寄与が大きくなり、観測されにくい単位が対象集団において代表する割合を推定に反映することができる。

以上の数式から分かるように、重み w_i は被験者 i の残差分散のみを変化させるものではない。被験者 i の平均構造、分散共分散構造、および観測値と予測値との乖離を含む、被験者 i の対数尤度への寄与全体に対して乗じられる。この点は、PROC MIXED の WEIGHT ステートメントによる精度重み付けと本質的に異なる。PROC MIXED の WEIGHT は、残差分散共分散行列を重みに応じて変換することにより、観測値の測定精度または異分散性をモデル化する。一方、尤度重みでは、被験者が対象集団において代表する程度を反映するため、被験者全体の尤度寄与を直接調整する。ここで、精度重みづけと尤度重みづけを整理する意味で、両者の違いを表 1 にまとめる。

表 1. 精度重みづけと尤度重みづけの違い

項目	精度重みづけ (Precision Weighting)	尤度重みづけ (Likelihood Weighting)
考え方	誤差分散の調整	頻度・代表性(寄与度)の調整
意味づけ	測定精度が高いデータほど分散が小さくなる。	重みが大きいほど、その観測値が対象集団においてより多くの頻度・重要度(代表性)を持つ。

2.3 SAS での実装コードの比較

PROC MIXED における精度重み付き MMRM の実装は、以下のコードのように WEIGHT ステートメントを単純に追加する形で行う。本コードでは、解析データセット targ を入力データとし、CHG を目的変数としている。CHG は、ベースラインからの変化量を表す連続変数とする。モデルには、ベースライン値 BASE、投与群 TRT、評価時点 AVISIT、および投与群と評価時点の交互作用 TRT*AVISIT を含めている。

```
proc mixed data=targ;
  class USUBJID TRT AVISIT;
  model CHG = BASE TRT AVISIT TRT*AVISIT / ddfm=kr solution;
  repeated AVISIT / subject=USUBJID type=un;
  weight WT; /* 精度重みとして WT を適用 */
  lsmeans TRT*AVISIT / diff cl;
run;
```

一方、PROC GLIMMIX では、重みを指定するオプションとして、MODEL ステートメントの OBSWEIGHT オプション（観測レベル重み）と、RANDOM ステートメントの WEIGHT オプション（被験者クラターレベル重み）が存在する。RANDOM ステートメントの WEIGHT オプションは、地域、施設、世帯、個人など、複数の階層において抽出確率が異なる場合を想定している。臨床試験の経時データ解析において、個々の被験者の代表性を補正する目的では、通常は前者の OBSWEIGHT を使用する。本コードでは、dist=normal を指定することで、連続アウトカムに対する正規分布モデルとしての解析を想定している。次に、プロシジャ名の後に指定した empirical オプションの指定によって、サンドイッチ型分散共分散行列を用いている。この指定は主として、固定効果の標準誤差、信頼区間および検定統計量の計算に影響するものであり、

固定効果の点推定値そのものを変更するものではない。また、RANDOM _RESIDUAL_ は、残差側（R-side）の分散共分散構造を指定するための記述である。subject=USUBJID により、同一被験者内の反復測定値間に相関があることをモデル化する。type=un は非構造化分散共分散構造を表す。

```
proc glimmix data=targ empirical;
 class USUBJID TRT AVISIT;
 model CHG = BASE TRT AVISIT TRT*AVISIT / dist=normal obsweight=WT;
 random _residual_ / subject=USUBJID type=un;
run;
```

ただし、PROC GLIMMIX では、観測値または被験者に固有の重みを OBSWEIGHT= 等で指定した場合、これらの重みを用いた解析では尤度ベースの推定法の指定が必要となるため、Gauss-Hermite adaptive quadrature を用いる METHOD=QUAD (Pfeffermann et al. (1998), Rabe-Hesketh and Skrondal (2006))等を採用し、モデルパラメータを最尤推定する必要がある。一方、METHOD=QUAD では、上記の RANDOM _RESIDUAL_ による R-side の残差分散共分散構造をサポートしない。したがって、本コードは、観測値レベルの尤度重みと R-side の非構造化共分散構造を同時に指定することを意図したコード例ではあるが、PROC GLIMMIX 内ではこの組合せを適合できない点に注意を要する。詳細は 3 章にて説明する。

3. 実務上の注意点

3.1 SAS PROC GLIMMIX を用いた実装上での注意点

個体別の尤度重みを考慮できる PROC GLIMMIX の使用を検討したが、2.3 章にて述べた通り、実務上の制約が生じる。経時データの繰り返し測定における時間相関構造を適切に考慮するため、R-side（残差側）の共分散モデル（random _residual_ / subject=USUBJID type=un; など）を指定し、かつ個体別重みを考慮するために MODEL ステートメントで obsweight=WT を指定すると、以下のエラーが発生する。

> ERROR: Subject or observation specific weights are only supported for nested models that are estimated using METHOD=QUAD.

エラーを回避するために推定手法を変更（例えば METHOD=QUAD 等に指定）すると、次に以下のエラーが発生する。

> ERROR: R-side random effects are not supported for METHOD=QUAD.

この挙動の背景には以下の数理的・システム上の制約があることが判明した。
1. GLIMMIX プロシジャにおいて、観測値または被験者個別の重み（尤度重み）を適用したモデルの適合は、推定手法として METHOD=QUAD を使用する場合にのみサポートされる。
2. 一方、METHOD=QUAD では、繰り返し測定の相関構造を表現する R-side（残差側）のランダム効果（UN、AR(1)、TOEP などの共分散構造）をモデリングすることができない。
3. したがって、OBSWEIGHT=による観測値・被験者固有の重みと、METHOD=QUAD 下での R-side 残差共分散構造を同時に適合することはできなかった。

この制約は、尤度重みを用いた経時データ解析の実装を検討する際に重要な留意点となる。

3.2 実務におけるその他の注意事項

上記の制約に対しては、単に利用可能なプロシジャを選択するのではなく、推定したい対象集団、重みの統計学的意味、および被験者内相関の表現を踏まえて、解析モデルと推測方法を一体として検討する必要がある。以下に、実務上の主な選択肢と解析計画上の留意点を示す。

- **1. G-side のランダム効果による代替（ランダム切片モデルの適用）:** R-side のランダム効果でモデル化することができないため、代わりに G-side のランダム効果を用いて被験者内相関をモデル化する。具体的には、以下のようにランダム切片モデルを適合させる。

```
proc glimmix method=quad empirical=classical;
  class USUBJID TRT AVISIT;
  model CHG = TRT AVISIT TRT*AVISIT / dist=normal obsweight=WT;
  random int / subject=USUBJID;
run;
```

  上記のモデルは、同一被験者内の異なる時点間の相関が一様であると仮定する。時間間隔に依存する相関（測定時点が近いほど相関が高く、離れるほど低い）はモデル化できないが、時点が異なっても相関がほぼ一定に近い状況であれば有効なモデルとなる。また、連続変数としての時点（AVISITN）を用いて random int AVISITN / subject=USUBJID type=un; としてランダム切片＋ランダム傾きモデルを当てることで、近い時点ほど相関が高く離れるほど低いという経時データの挙動をある程度近似して表現できる場合がある。ただし、誘導される相関構造はランダム効果の分散共分散構造により制約され、UN や AR(1)等の R-side 共分散構造を一般に再現するものではない。そのため、対象データに対する近似の妥当性は、適合度および感度分析等を通じて確認する必要がある。
- **2. PROC NLMIXED による多変量正規尤度の直接実装:** R-side の UN（非構造化）共分散行列等を数学的に維持しつつ、各被験者に対して厳密な尤度重みを掛けたい場合は、PROC GLIMMIX ではなく PROC NLMIXED を用いることも可能である。NLMIXED 内で経時データの多変量正規尤度を直接数式として定義し、その対数尤度に重み WT を乗じた上で最尤法によるパラメータ探索を行う。プログラミングの難易度は高いが、数理的整合性を最優先とする場合の解決策として挙げられる。
- **3. 臨床試験の意思決定・当局相談における留意点:** 臨床試験の統計解析計画書（SAP）を作成する際には、使用する SAS のプロシジャ制限を正確に把握し、解析モデルの選択（R-side の共分散構造を諦め G-side ランダム切片とするか、あるいは NLMIXED による直接プログラミングとするか）について、あらかじめ規制当局や臨床統計の専門家と十分な合意形成を図ることが必須である。

4. シミュレーション

PROC MIXED の WEIGHT 指定と NLMIXED による尤度重み付けとの間に、どれくらいの治療効果の点推定値に関する実質的な差が見られるのかシミュレーションを実施した。本シミュレーションでは、対象集団への選択確率が共変量 X に依存する状況を設定し、被験者の代表性を補正する重みを用いた場合の推定性能を評価した。具体的には、以下の 3 手法を比較した。

・重み無し MMRM（PROC MIXED）

・PROC MIXED の WEIGHT ステートメントに指定した MMRM

・被験者単位で対数尤度に重み(IPW)を乗じた尤度重み付き多変量正規モデル（PROC NLMIXED）

各反復における対象集団の被験者数は 1,500 例とし、二値共変量 X は確率 0.5 で生成した。解析対象への選択確率は、X=0 の被験者では 0.80、X=1 の被験者では 0.20 とした。このため、選択後の解析集団では X=1 の被験者が過少となる設定とした。アウトカムは 4 時点の反復測定値とし、被験者内相関を有する多変量正規分布に従う誤差を発生させた。最終時点における治療群差は、X=0 で−2.0、X=1 で−3.0 とした。対象集団における X=1 の割合は 0.5 であることから、対象集団における最終時点の平均治療群差の真値は、条件付き治療群差の期待値として、(−2.0)*0.5 + (−3.0)*0.5 ＝−2.5 とした。各手法について 500 反復のシミュレーションを実施し、最終時点における治療群差の平均推定値、バイアス、モンテカルロ標準偏差（Monte Carlo standard deviation; MC SD）、平均モデル標準誤差を評価した。ここでのバイアスは、平均推定値−真値として算出を行った。それぞれの算出結果を表 2 に示す。

表 2: 最終時点の治療群差に関するシミュレーション結果

解析方法	平均推定値	バイアス	MC SD	平均モデル標準誤差
PROC MIXED WEIGHT=IPW	−2.501104	-0.00110	0.090996	0.076112
重み無し MMRM	−2.199045	0.30095	0.072194	0.074783
尤度重み付きモデル(NLMIXED)	−2.501096	-0.00109	0.090998	0.075569

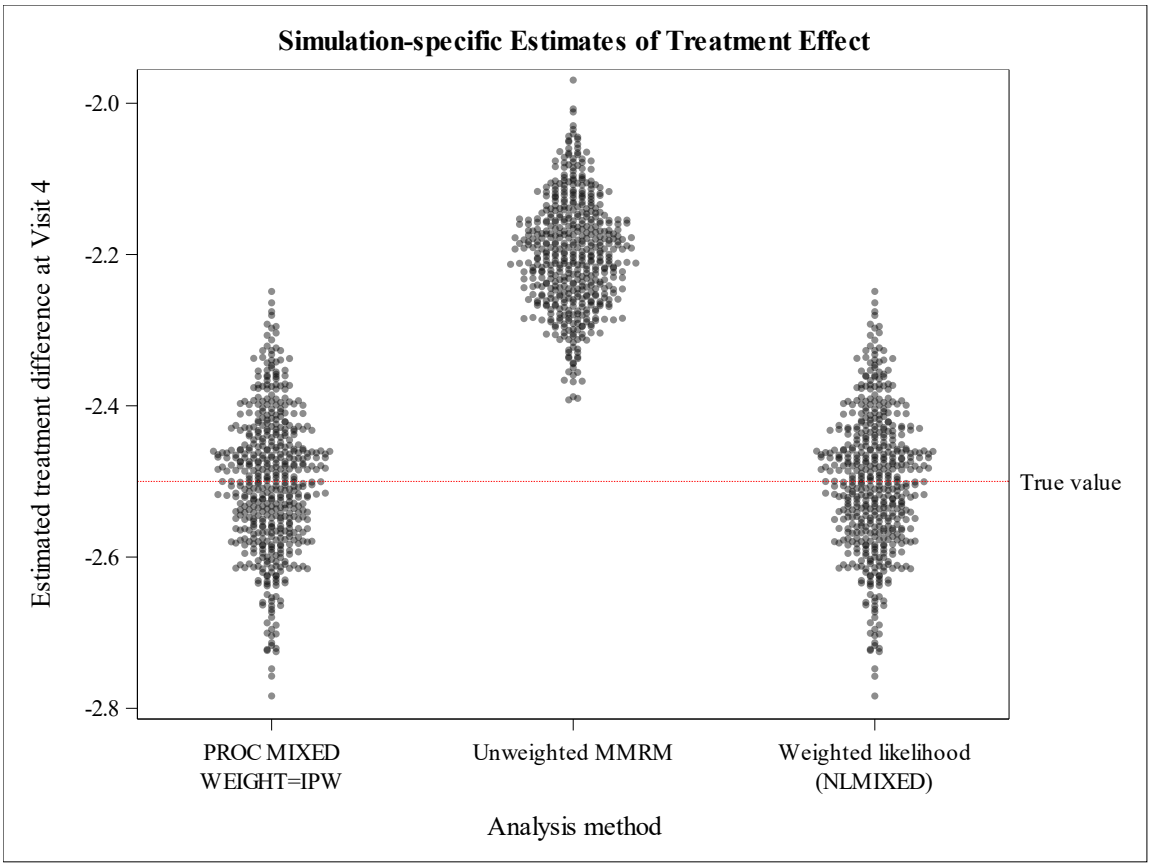


図 1. 各反復における最終時点の治療群差推定値

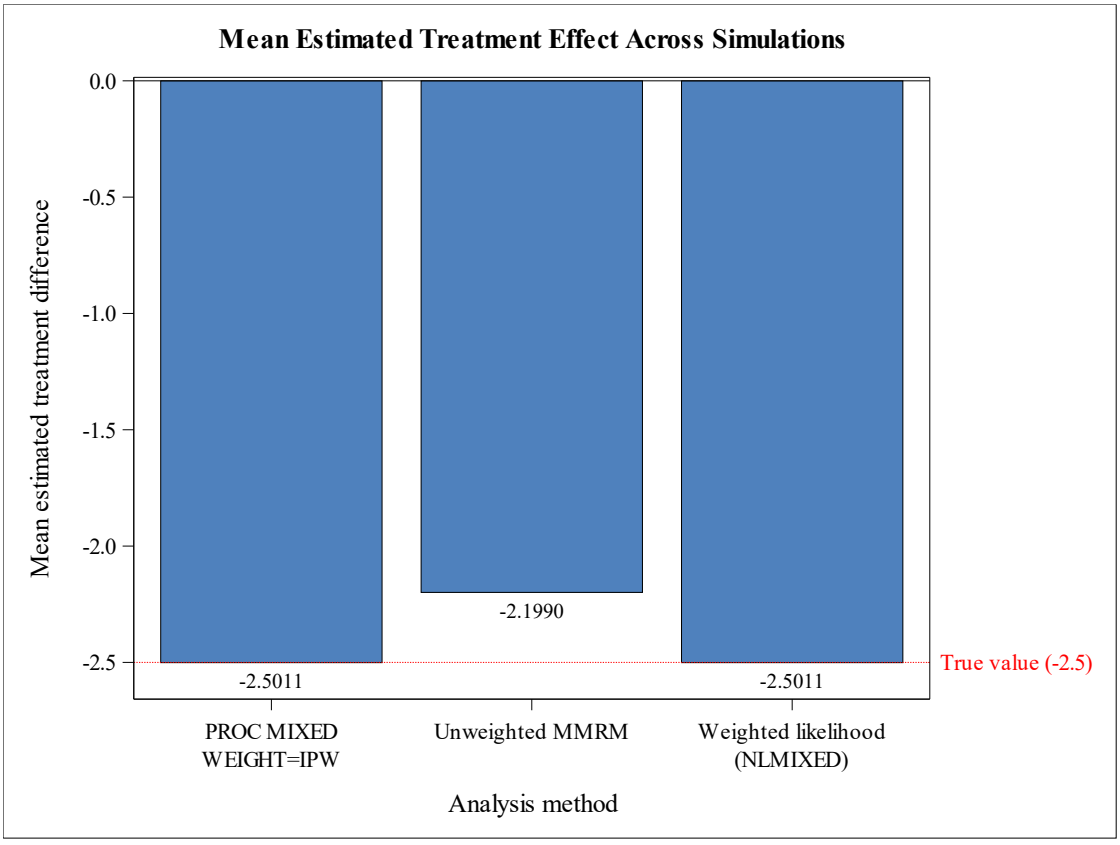


図 2. 最終時点の治療群差推定値の分布

図 1 に、各反復における最終時点の治療群差推定値を示す。重み無し MMRM による推定値は、真値である−2.5 より一貫して大きい値、すなわち治療効果を過小評価する方向に分布した。一方、PROC MIXED の WEIGHT ステートメントを使用した方法および NLMIXED による尤度重み付きモデルでは、推定値の分布の中心はともに真値付近に位置した。図 2 に、治療群差推定値の分布を示す。重み無し MMRM の平均推定値は−2.1990 であり、真値−2.5 に対するバイアスは約 0.3010 であった。これは、選択後の解析集団では X=1 の被験者が過少となり、当該集団における平均治療効果が対象集団における平均治療効果より小さくなるように設定されていたことと整合的である。これに対し、PROC MIXED の WEIGHT ステートメントに IPW を指定した方法および NLMIXED による尤度重み付きモデルの平均推定値はいずれも−2.5011 であり、いずれも真値−2.5 との差は約−0.0011 であり、実質的なバイアスは認められなかった。

また、PROC MIXED の WEIGHT=IPW および尤度重み付き NLMIXED の結果は、平均推定値だけでなく、バイアス、MC SD および平均モデル SE についても近い値を示した。具体的には、MC SD はそれぞれ 0.0910、平均モデル SE はそれぞれ 0.0761 および 0.0756 であった。ただし、PROC MIXED の WEIGHT ステートメントによる重み付けと、NLMIXED で実装した尤度重み付けは、数理的解釈としては異なる。本シミュレーションにおいて両者の平均治療効果推定値が一致したことは、PROC MIXED の WEIGHT ステートメントが尤度重みを一般に再現することを示すものではない。今回の設定では、以下の特徴があった。

・重みは被験者単位で定義され、同一被験者の全評価時点に共通の値であった。

・全被験者について 4 時点のアウトカムが観測されていた。

このような比較的単純で均衡的な条件下では、PROC MIXED における精度重み付けであっても、平均構造の推定に関しては被験者単位の重み付き推定と同様の挙動を示し得る。その結果、最終時点の治療群差に関する点推定値が NLMIXED による尤度重み付き推定値と一致したものと考えられる。欠測パターンが被験者間で異なる場合、重みが時点ごとに異なる場合、平均構造に時点に関する制約を置く場合、あるいはより複雑なランダム効果構造を用いる場合には、点推定値にも差が現れる可能性がある。したがって、PROC MIXED の WEIGHT ステートメントに重みを指定して本シミュレーションと同様の結果が得られたとしても、これを、被験者代表性を補正する尤度重み付き MMRM として一般化して解釈することには注意を要する。実務では、重みの作成目的、推定したい対象集団、モデルの平均・共分散構造、および標準誤差の推定方法を明確にした上で、使用するプロシジャおよび実装方法を選択する必要がある。

5. まとめ

本稿では、反復測定混合モデルにおける重みの取り扱いについて、SAS PROC MIXED の WEIGHT ステートメントによる重み付けと、被験者または観測値の対数尤度への寄与に直接重みを乗じる尤度重み付けとの数理的・実装上の相違を整理した。

PROC MIXED の WEIGHT ステートメントは、観測値ごとの重みを測定精度、すなわち残差分散の逆数として扱う精度重み付けである。したがって、重みの大きい観測値は残差分散が小さい、より精度の高い観測値としてモデル化される。一方、逆確率重みなど、被験者が対象集団において代表する程度を反映するため

の重みは、各被験者または観測値の尤度寄与を調整する尤度重みとして解釈される。このように、両者は重みが作用する対象および統計学的な意味が本質的に異なる。

また、PROC GLIMMIX では、OBSWEIGHT=等を用いて観測値または被験者に固有の尤度重みを考慮する実装が考えられる。しかし、個別重みを用いる場合には METHOD=QUAD が必要となる一方、METHOD=QUAD では RANDOM _RESIDUAL_による R-side の残差分散共分散構造を指定できない。このため、非構造化共分散構造（UN）等を用いた通常の MMRM と、個別の尤度重みを PROC GLIMMIX 内で同時に実装することには制約があることを示した。

シミュレーションでは、選択確率が共変量 X に依存し、選択後の解析集団において X=1 の被験者が過少となる状況を設定した。重み無し MMRM では、対象集団における最終時点の真の平均治療群差である−2.5 に対して、平均推定値は−2.1990 であった。一方、IPW を PROC MIXED の WEIGHT ステートメントに指定した解析および NLMIXED で被験者単位の尤度重みを実装した解析では、平均推定値はいずれも−2.5011 であり、真値に極めて近かった。ただし、本シミュレーションにおいて PROC MIXED の WEIGHT 指定と尤度重み付き NLMIXED の点推定値が一致したことは、両者の重み付けが同一の意味を持つこと、あるいは PROC MIXED の WEIGHT が一般に尤度重みを代替できることを意味するものではない。今回の設定では、重みが被験者単位で定義されて同一被験者の全時点に共通であったこと、全被験者について 4 時点の測定値が得られていたこと、および比較的均衡的な平均構造を用いたことが、両手法の点推定値が近接した要因と考えられる。実データでは、脱落や間欠欠測により被験者ごとに観測時点が異なる場合、時点ごとに異なる重みを用いる場合、重みがアウトカムや欠測パターンと複雑に関連する場合、あるいは平均構造や共分散構造に制約を置く場合がある。このような状況では、精度重みと尤度重みの相違が、治療効果の点推定値のみならず、標準誤差、信頼区間および検定結果にも影響する可能性がある。したがって、PROC MIXED の WEIGHT ステートメントを指定する場合には、単に「重み付き解析を実施した」と解釈するのではなく、その重みが測定精度を表すものなのか、対象集団に対する代表性を補正するものなのかを明確にする必要がある。

対象集団への一般化、選択バイアスの補正または欠測・脱落に対する重みづけ解析を目的とする場合には、推定対象（estimand）、重みの定義、平均構造、被験者内相関の表現、および標準誤差の算出方法を一体として検討することが重要である。R-side の柔軟な共分散構造を維持しつつ尤度重みを厳密に適用する必要がある場合には、PROC NLMIXED による多変量正規尤度の直接実装が選択肢となる。ただし、この方法ではプログラムの実装および推測の妥当性確認がより重要となる。実務においては、解析目的と重みの統計学的意味に適合した方法を選択し、必要に応じてシミュレーション、ロバスト分散推定またはリサンプリング法等によって推定および推測の性能を確認することが望まれる。

6. 参考文献

1. SAS Institute Inc. (2025). "The MIXED Procedure: WEIGHT Statement". SAS/STAT User's Guide. https://documentation.sas.com/doc/zh-TW/pgmsascdc/v_041/statug/statug_mixed_syntax17.htm
2. SAS Institute Inc. (2025). "The GLIMMIX Procedure: Weighted Multilevel Model for Survey Data". SAS/STAT User's Guide. https://documentation.sas.com/doc/en/pgmsascdc/v_074/statug/statug_glimmix_details25.htm
3. SAS Institute Inc. (2025). "The GLIMMIX Procedure: Pseudo-likelihood Estimation for Weighted Multilevel Models". SAS/STAT User's Guide. https://documentation.sas.com/doc/en/pgmsascdc/v_074/statug/statug_glimmix_details02.htm

4. Pfeffermann, D., Skinner, C. J., Holmes, D. J., Goldstein, H., and Rasbash, J. (1998). “Weighting for Unequal Selection Probabilities in Multilevel Models.” Journal of the Royal Statistical Society, Series B 60:23–40.

5. Rabe-Hesketh, S., and Skrondal, A. (2006). “Multilevel Modelling of Complex Survey Data.” Journal of the Royal Statistical Society, Series A 169:805–827.