

## 論文の内容の要旨

論文題目 縦断的データに経時横断的データを加えた場合のベイズ推測

氏名 石井秀宗

教育科学の分野においては、検査の再検査信頼性係数を推定する場合や、成長や発達による特性の変化、トレーニングや教育の効果などについて研究する場合など、繰り返し測定を行い縦断的データを収集して研究を進めることがある。しかし、実際の研究では、測定の各時点ではある程度の被験者数を確保することができるが、全ての時点で観測される被験者はそれほど多くはないという状況、つまり、経時的に（各時点ごとに）観測される横断的データ（本論文では「経時横断的データ」または単に「横断的データ」と記述した）はある程度収集することができるが、縦断的データは少数しか収集できないという状況が生じる場合がある。本論文は、教育科学の分野において縦断的研究を行う際に発生するこのような問題に鑑み、縦断的データに経時横断的データを加えた場合の統計的推測を行うモデルを作成し、母集団を規定するパラメタの推定において、経時横断的データを加えることの効果を解析的に考察することを目的とした。統計的手法としてはベイズ推定法を用い、パラメタの事前分布には公的事前分布、同一母集団からの標本のモデル分布には母分散が等質な（多変量）正規分布を設定した。

第1章「序論」では、縦断的データと通常の横断的データの違いについてま

ず説明した。そして、縦断的データの最も単純な場合である 2 回測定データに注目して、教育科学の分野でもよく用いられる再検査信頼性係数の推定、および、対応のある 2 群の平均値差の検定・推定の方法について解説した。一般的な縦断的データを扱う統計的手法の概説も行った。また、縦断的データを扱う上で問題となる欠測に関して、欠測の種類と、従来の統計的手法における欠測の取り扱い方について言及し、縦断的データに積極的に横断的データを加えて分析するという本論文における発想が、従来の欠測という消極的な考えとは一線を画すものであることを述べた。

第 2 章「ベイズ推定法」では、本論文で統計的手法として用いたベイズ推定法について解説した。そして、ベイズ推定法の持ついくつかの特質が、母集団を規定するパラメタの推定において横断的データを加えることの効果を解析的に考察するという本論文の目的によく適合するものであることを述べた。ベイズ推定法では確率モデルの設定の仕方が重要な意味を持つことから、事後予測  $p$  値によるモデルチェックの方法についても説明した。

第 3 章「2 回測定データの分析」では、教育科学の研究でもよく収集される 2 回測定データについて、縦断的データだけの場合と、横断的データを加えた場合の母平均、母分散、母相関係数の推定に関する考察を行った。

再検査信頼性係数の推定に対応させた母平均が共通なモデルでは、母平均パラメタの条件付き周辺事後分布として一般化  $t$  分布、母分散パラメタの条件付き周辺事後分布として逆  $\chi^2$  分布が導出された。母平均の差の推定に対応させた母平均が異なるモデルでは、母平均パラメタの条件付き周辺事後分布として 2 変量  $t$  分布、母平均パラメタの差の条件付き周辺事後分布として一般化  $t$  分布、母分散パラメタの条件付き周辺事後分布として逆  $\chi^2$  分布が導出された。そして、いずれの分布においても、横断的データを加えることにより事後分散が小さくなり、推定精度が向上することが明らかにされた。

母相関係数パラメタについては、母平均が共通なモデルと母平均が異なるモデルの双方において、縦断的データだけの場合と、横断的データを加えた場合に付け加え、横断的データの標本の大きさが大きい場合の周辺事後分布を導出した。そして、標本平均または標本分散の均一性が失われるほど母相関係数パラメタの事後モードが小さくなることが確認された。母平均が共通なモデルに従って考えるとこれは、再検査信頼性係数の推定において、平均や分散などの標本統計量の値が 2 回の測定で不均一であるほど、再検査信頼性係数が小さく

推定されることを意味するものであると理解された。

また、縦断的データだけの場合と同様に、横断的データの標本の大きさが大きい場合においても、フィッシャーの  $z$  変換を用いて母相関係数パラメータを変換したものの近似事後分布として正規分布が導かれることが明らかにされた。近似事後分布の分散の大きさを比較したところでは、標本統計量の値がある程度均一であれば、横断的データを加えることにより事後分散が小さくなり、推定精度が向上することが明らかにされた。そして、縦断的データの標本の大きさが小さいほど、また、標本相関係数の値が大きいほど、推定精度の向上の割合は大きいことなどが確認された。近似事後分布に関するこれらの結果から、母相関係数パラメータについても、横断的データを加えることにより事後分散が小さくなり、推定精度が向上することが明らかにされた。

第4章「任意回繰り返し測定を行ったデータの分析」では、測定の繰り返し数を任意回に一般化した複数回測定モデル、および、被験者間要因のある複数回測定モデルを作成し、第3章と同様の考察を行った。測定時点や被験者間要因の主効果、および、それらの交互作用を検討する母平均パラメータの線形結合についても検討した。

複数回測定モデルでは、母平均パラメータ、および、母平均パラメータの線形結合の条件付き周辺事後分布としてそれぞれ多変量  $t$  分布、母分散パラメータの条件付き周辺事後分布として逆  $\chi^2$  分布が導出され、いずれの分布においても、横断的データを加えることにより事後分散が小さくなり、推定精度が向上することが明らかにされた。

母相関係数パラメータについては、一般の母相関係数行列の周辺事後分布を導出するとともに、級内相関パターン行列の母相関係数パラメータの周辺事後分布を導出した。そして、横断的データを加えることにより事後分散が小さくなり、推定精度が向上することが確認された。

被験者間要因のある複数回測定モデルについては、1) 母分散はすべての変数を通して共通、母相関係数行列も被験者間要因の水準を超えて共通とする、2) 母分散は被験者間要因の水準内で共通、母相関係数行列は被験者間要因の水準を超えて共通とする、3) 母分散は被験者間要因の水準内で共通、母相関係数行列は被験者間要因の各水準ごとに設定する、という3つのモデルを作成した。

その結果、母平均パラメータの条件付き周辺事後分布として、1) では多変量  $t$  分布、2), 3) では被験者間要因の各水準ごとに定まる多変量  $t$  分布の積が導出された。また、母平均パラメータの線形結合の条件付き周辺事後分布として、1) で

は多変量  $t$  分布, 2),3)では被験者間要因の各水準ごとに定まる多変量  $t$  分布の積が導出された. さらに, 2),3)の母平均パラメタ, および, 母平均パラメタの線形結合の条件付き近似周辺事後分布として, それぞれ多変量正規分布が導出された. 母分散パラメタについては, 各母分散パラメタの条件付き周辺事後分布として逆  $\chi^2$  分布が導出された. いずれの分布においても, 横断的データを加えることにより事後分散が小さくなり, 推定精度が向上することが明らかにされた.

母相関係数パラメタについては, 各母相関係数行列の周辺事後分布を導出するとともに, 級内相関パターン行列の母相関係数パラメタの周辺事後分布を導出した. そして, 横断的データを加えることにより事後分散が小さくなり, 推定精度が向上しうることが確認された.

第5章「実データを用いた分析例」では, 第3章, 第4章で作成したモデルを用いて実際のデータを分析した例を示した. そして, 縦断的データだけの場合よりも横断的データを加えた場合の方が精度のよい推定になることが確認された. 母分散パラメタや母相関係数パラメタの設定が異なる複数のモデルで同一のデータを分析したところ, 母平均パラメタや母平均パラメタの線形結合の推定については, いずれのモデルを用いてもほぼ同様の結果になることが示された. また, 推定すべきパラメタの数が多いモデルほど, 全体の推定精度は低下することも確認された. なお, 事後予測  $p$  値によるモデルチェックの結果から, いずれの分析例においても, モデルが不適切であると積極的に考える必要はないことが確認された.

第6章「まとめと今後の展望」では, 本論文のまとめと今後の展望について述べた. 本論文では, 縦断的研究をする際に, 縦断的データに経時横断的データを加えて分析することを考え, その分析モデルとして, 2回測定モデル, 複数回測定モデル, 被験者間要因のある複数回測定モデルを作成した. そして, 母平均, 母分散, 母相関係数などの推定精度が, 経時横断的データを加えることにより向上することを解析的に明らかにした.

今後の展望としては, 部分的な縦断的データを加えた場合の分析や, 繰り返し測定の過程で縦断的に測定される被験者が順次交代していく場合の分析などについて検討することが考えられることを述べた. また, 母分散の等質性の仮定を緩和したモデルなど, 数値シミュレーション的な検討を必要とする研究の方向性もあることを提示した.