

経時測定データに対する関数データ解析 (Functional Data Analysis; FDA) の活用

経時測定データから曝露あるいは介入の効果を推定したい場合、同一個人に由来するデータ間の非独立性が問題となる。経時測定データへの解析手法としては、非独立性を考慮せず、各地点を別個の別個の変数として扱うモデルや、データ間の非独立性を考慮した混合モデルなど数多く存在するが、いずれの手法でも対処が不可能なケースが 2 つ考えられる。

1 つ目が、経時測定データが非常に多くの観測地点を持つケースである。観測地点の増加はそのままモデルの次元数の増加につながり、推定が安定しない可能性がある。

2 つ目が、データが非単調な欠測を伴うケースである。経時測定データが非単調な欠測をもつ場合の解析方法は確立されていない。また、欠測は観測時点あるいは時点数の不均一さをもたらす可能性があり、特に時点の不均一さが存在した場合には古典的な多変量解析手法を適用することは困難となる。

この 2 点に対応可能な手法として、複数時点のデータを時間の関数へと変換し、推定された関数をデータとして扱う関数データ解析 (Functional Data Analysis; FDA) が研究されている。関数化はデータの高次元化を抑制し、連続時間上でのデータの表現により個体間の時点・時点数の不均一さを解消する。スカラーデータに対する統計解析手法の FDA への拡張も数多く提案されており、いくつかの論文では既存の予測モデルに対する予測精度の向上なども報告されている。

本抄読会では、垂直跳び中の膝関節運動に対し FDA を適用した Ryan らの研究を例に取り、変数の関数化および関数化したデータへの関数主成分分析 (Functional PCA; FPCA) の適用方法について紹介する。

参考文献

- 1 J.O.Ramsay and B.W.Silverman. Functional Data Analysis. Springer, 2005.
- 2 松井秀俊「関数データに基づく統計的モデリング」統計数理 (2019) 第 67 巻 第 1 号 73 - 96
- 3 Sørensen H, Goldsmith J and Sangalli LM. An introduction with medical applications to functional data analysis. Stat Med. 2013 Dec 30;32(30):5222-40. doi: 10.1002/sim.5989.
- 4 Willie Ryan, Andrew Harrison and Kevin Hayes. Functional data analysis of knee joint kinematics in the vertical jump. Sports Biomech. 2006 Jan;5(1):121-38. doi: 10.1080/14763141.2006.9628228.
- 5 阿部貴行「欠測データの統計解析」朝倉書店 (2016)