

区間打ち切りデータに対する生存関数推定における 複数の補助変数を用いたアプローチ

特定の事象が発現するまでの時間に興味があるような医学研究においては、区間打ち切りデータが発生することがある。例えば、X線測定に基づく骨粗鬆症性の椎体骨折発現までの時間を比較するような医学研究では、X線の測定は来院ごとに行われるため、次の来院時に骨折が判明するような場合は区間打ち切りデータの一例となる。

区間打ち切りデータの解析手法については、広くコンセンサスの取れた手法がないこともあり、区間の midpoint (もしくは、始点や終点) をイベント発現タイミングと見なし、右側打ち切りデータの解析手法として提案された手法を便宜的に当てはめることなどが行われるが、これらのアプローチは妥当な解析手法ではない事が知られている¹⁾。

区間打ち切りデータにおける生存関数の推定方法としてはいくつかの提案があるものの^{2),3)}、これらはイベント発現時間のみを推定に用いる。一方、イベント発現時点の情報に加え、発現タイミングに関する情報が得られる場合があり、このようなデータの下ではこの情報を推定に利用することにより、より正確な生存関数が推定されることが期待される。

前回の抄読会では、区間打ち切りデータにおけるイベント発現時点に関連する補助変数が複数存在する場合に、これらを利用した生存関数推定における提案法について紹介した。本抄読会では提案法のシミュレーション実験の結果及び、実データへの適用結果について述べる。

参考文献

- 1) Lindsey JC, Ryan LM. Tutorial in Biostatistics: Methods for interval-censored data. *Statistics in Medicine* 1998; **17**:219-238.
- 2) Peto R. Experimental survival curves for interval-censored data. *Applied Statistics* 1973; **22**:86-91.
- 3) Turnbull BW. The empirical distribution function with arbitrarily grouped, censored, and truncated data. *Journal of the Royal Statistical Society, Series B* 1976; **38**:290-295.