

論文内容の要旨

論文題目 構造ネスト平均モデルの因果パラメータに対する疑似尤度法に基づく推測

指導教員 松山 裕 准教授

東京大学大学院医学系研究科

平成 19 年 4 月進学

博士後期課程

健康科学・看護学専攻

田栗 正隆

1 背景と目的

ランダム割り付け比較試験は人間に対して倫理的に許される範囲で実施される科学的な実験であり、治療の科学的効果を知るために最も重要な研究デザインである。しかし、全ての対象者が最初に割り付けられた治療を追跡が終了するまで必ずしも遵守し続けるとは限らず、対象者によっては割り付けられなかった方の治療を受けることもある。このようなデータに対して標準的に用いられている解析方針として、ランダム割り付け後の治療経過によらず、割り付けの対象となった全ての対象者を割り付けられた治療グループに属するとして解析する Intention-To-Treat の原則に基づく解析 (ITT 解析) が挙げられる。しかしながら、ITT 解析では、割り付けられた治療を守らなかった対象者を元々の群として扱うため、ノンコンプライアンス割合が高くなるにつれて、もし試験に参加した対象者全員が割り付け治療を遵守した場合に観測される治療効果 (因果効果) を大きく過小評価しうる。

ランダムでないノンコンプライアンスが存在する状況において因果治療効果を推定する方法論として、近年 Rubin の因果モデルに基づく統計的方法が発展してきた。Robins が提案した構造ネスト平均モデルでは、観測された結果変数と、基準となる「試験治療を受けていない」場合の結果変数の差 (あるいは比) の期待値を、未知パラメータを用いてモデル化する。Robins は、「試験治療を受けていない」場合の結果変数 (潜在結果変数) が性・年齢などと同様の治療前変数として考えることができるため、治療割り付けとは独立になることから、構造ネスト平均モデルの未知パラメータの推定を行うことを提案している (G 推定法)。Robins はさらに結果変数が経時的に測定されている場合への一般化も行って

いる。

結果変数が経時的に繰り返し測定されている場合には、治療と時間の交互作用や治療とベースライン共変量との交互作用に興味がある場合がある。この交互作用効果に関する問題は、通常の回帰モデルにおける変数選択の問題と同様に、どのような構造モデルがデータにより当てはまっているか、というモデル選択の問題を喚起する。しかしながら、現在のところ構造ネスト平均モデルにおけるモデル選択基準は提案されておらず、未知の因果パラメータが複数存在するような構造モデルを推測の候補として考える場合、構造ネスト平均モデルに対する統計学的方法論は、十分存在するとは言い難い状況にある。

そこで本論文では、構造ネスト平均モデルに対する G 推定量が推定関数に基づく M 推定量に属することを利用して、G 推定関数に対する疑似尤度関数を導き、その疑似尤度関数に基づいて種々の統計的推測を行うことを提案した。

2 提案する疑似尤度法に基づく推測

Robins が提案した G 推定量の推定関数 $\psi(\theta)$ を用いて、その目的関数である疑似尤度関数 $\rho(\theta)$ を以下のように定義した。

$$\rho(\theta) = \int_s^\theta \psi(t) dt$$

ここで θ は関心のある治療効果パラメータ（ベクトル）であり、定義より $\psi(\theta) = \text{カ}\rho(\theta)/\text{カ}\theta$ を満たす。また、積分の下端 s は $\rho(0) = 0$ を満たすように決定する。上の積分（線積分）の結果は一般に経路に依存するが、本研究では積分結果が一意に定まるような $\psi(\theta)$ を発見し、それに基づいて推測を行うことを提案した。また、ノンコンプライアンスの補正を行う場合、割り付けが治療に正の効果を与えているという現実的な状況では、 $\rho(\theta)$ が $\hat{\theta}$ で最大値を持つであろうことも確認した。

具体的に、提案した疑似尤度関数 $\rho(\theta)$ を用いて、以下の 3 種の漸近検定を行うことができる。治療効果を規定する興味のあるパラメータベクトル θ を $\theta^T = (\theta_{(1)}^T, \theta_{(2)}^T)$ と分割し、 $H_0: \theta^T = (\theta_{(1)}^T, 0^T)$ という帰無仮説の検定を考える。ここで $\theta_{(1)}$ 、 $\theta_{(2)}$ はそれぞれ v_1 次元、 v_2 次元のパラメータベクトルとする。ここで $v = v_1 + v_2$ である。同様な v 次元のベクトル b の分割を $b_{(i)}$ 、正方行列 B の分割行列を $B_{(ij)}$ ($i, j = 1, 2$) と表す。すると、

(i) ワールド検定

$$W_n^2 = \hat{\theta}_{(2)}^r \hat{V}(\theta)_{(22)}^{-1} \hat{\theta}_{(2)}$$

(ii) スコア検定

$$S_n^2 = \bar{\psi}_{(2)}(\hat{\theta}_0) \hat{C}(\theta) \bar{\psi}_{(2)}(\hat{\theta}_0)$$

(iii) 尤度比検定

$$L_n^2 = 2 \cdot \sum_{i=1}^n \{ \rho_i(\hat{\theta}) - \rho_i(\hat{\theta}_0) \}$$

をそれぞれ導くことができる。ここで $\hat{\theta}_0$ は帰無仮説の下での θ の推定量である。 W_n^2 、 S_n^2 はそれぞれ帰無仮説の下で漸近的に自由度 ν_2 のカイ 2 乗分布に従う。また、 L_n^2 は独立な自由度 1 のカイ 2 乗分布に従う確率変数の重みつき和に分布収束する。

また、これらの検定統計量を用いた信頼区間の構成も可能である。パラメータが 1 つのケースについて $100(1 - \alpha)\%$ 信頼区間を構成する方法について説明する。まず、ワールド型の信頼区間は

$$\hat{\theta} \mp Z_{1-\alpha/2} \text{se}(\hat{\theta})$$

により構成できる。ここで $Z_{1-\alpha/2}$ は標準正規分布の $100(1 - \alpha/2)\%$ 点である。次にスコア型の信頼区間は

$$\frac{\bar{\psi}(\theta)^2}{n^{-1} J_n(\theta)} \Big| \chi_{(1-\alpha)}^2$$

を満たす θ の集合として定義される。ここで $\chi_{(1-\alpha)}^2$ は自由度 1 のカイ 2 乗分布の $100(1 - \alpha)\%$ 点である。

最後に尤度比型の信頼区間は

$$\frac{2 \cdot \sum_{i=1}^n \{ \rho_i(\hat{\theta}) - \rho_i(\theta) \}}{\hat{\lambda}} \Big| \chi_{(1-\alpha)}^2$$

を満たす θ の集合として定義される。

さらに、導出した疑似尤度関数を利用して AIC と同様のモデル選択基準 GQIC (Quasi-likelihood based Information Criterion for G-estimation) を提案した。

$$\text{GQIC} = -2 \cdot \sum_{i=1}^n \rho_i(\hat{\theta}) + 2 \text{trace}(\hat{I}_n^{-1} \hat{J}_n)$$

GQIC はパラメータ θ_* で規定される真のモデルと、パラメータ θ で規定される仮定したモデルの疑似尤

度の距離 $\Delta_n(\hat{\theta}, \theta_*) = n \text{tr} E_M \left[\frac{1}{2} \rho(\hat{\theta}) \right]$ の良い推定量となっている。

3 シミュレーション研究

GQIC の性能評価を評価するためにシミュレーション研究を行った。真のモデルが治療の主効果（切片）のみである場合やベースライン共変量や時点との交互作用もある場合等、複数の組み合わせを想定したが、真のモデルが切片のみである場合を除いて、どのパターンにおいても GQIC は正しいモデルを選択する割合が最も高かった。特にサンプルサイズが 500 以下のときは、P 値による選択基準と比較して性能の差が顕著であった。結果の 1 部を表 1 に示す。提案する GQIC では 68.2%と検定の P 値に基づく選択結果 (Z-test1, Z-test2) と比較して、非常に高い割合で真のモデル（表 1 の 1 番下の行）を選択した。

表 1 シミュレーション結果(g(X)交互作用なし、n=250)

true θ	Model	Method			
		GQIC	GQIC2	Z-test1	Z-test2
(-3,-2,2,2)	(1)	0.067	0.085	0.041	0.031
	(2)	0.052	0.046	0.018	0.014
	(3)	0.181	0.205	0.028	0.016
	(4)	0.018	0.039	0.020	0.025
	(5)	<u>0.682</u>	<u>0.625</u>	<u>0.084</u>	<u>0.187</u>

数値は各モデルの選択割合を表す

Z-test1;5%水準,Z-test2;10%水準の P 値による選択

true θ は治療効果のモデルのパラメータ真値

4 MEGA Study データの解析

提案した方法の適用事例として、スタチン系薬剤であるプラバスタチンの冠動脈イベント発生の抑制効果を検証することを目的とした大規模一次予防試験である MEGA Study データのデータ解析を行った。MEGA Study の対象者は、本解析での最終追跡時点である 5 年時においては、試験治療群であるプラバスタチン群では 33.1%、対照群である食事療法群では 16.3%の対象者が割り付けを遵守していなかった。解析の目的は、コンプライアンスの影響を補正した LDL 値改善に対するプラバスタチンの因果効果を推定し、時点やベースライン共変量と治療の交互作用を考慮した解析を行うことである。ITT 解析と G 推定法 (G-est) により推定された群ごとの LDL 値の平均的な推移とその 95%信頼区間を図 1 に示す。

また、MEGA Study の全対象者の内のサブ集団をランダムに繰り返しサンプリングすることによって、提案した疑似尤度に基づくモデル選択基準の性能を様々なサンプルサイズにおいて確認した。その結果、提案する GQIC はどのサンプルサイズにおいても、最も高い割合で元データに 1 番良く当てはまったモデルを選択していた。なお、当てはめたモデルは、共変量として「モデル(1); LDL」、「モデル(2); LDL、高脂血症治療歴有り」、「モデル(3); LDL、高血圧有り」、「モデル(4); LDL、高脂血症治療歴有り、高血圧有り」、「モデル(5); LDL、高脂血症治療歴有り、高血圧有り、年齢」を含める 5 つのモデルである。結果の一部を表 2 に示す（全データに最も良く当てはまったモデルは、下から 2 行目）。

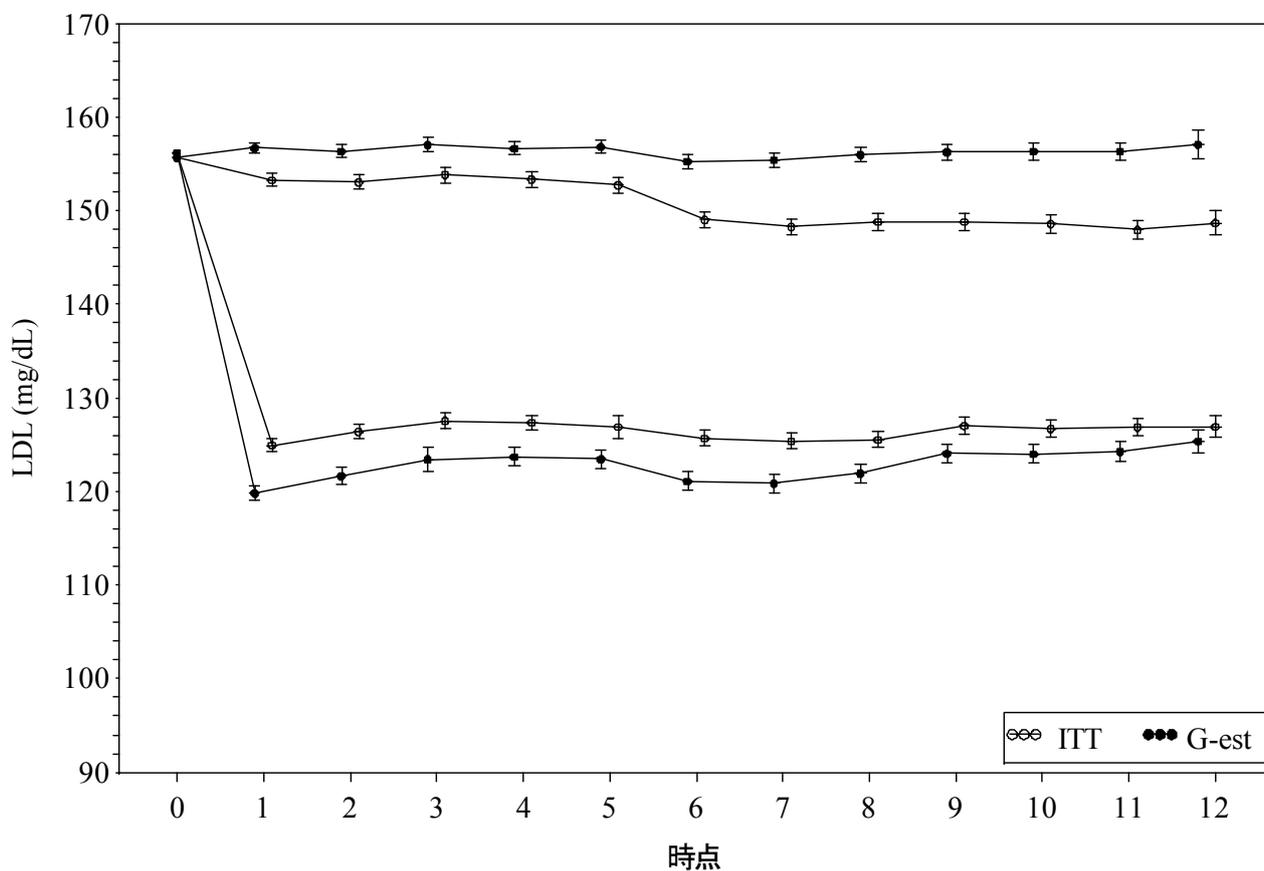


図 1 割り付け群ごとの LDL 値の平均的な推移とその 95%信頼区間

表 2 1,000 回の復元抽出による MEGA Study サブ集団におけるモデル選択結果

n	Model	Method			
		GQIC	GQIC2	Z-test1	Z-test2
5,000	(1)	0.028	0.000	0.001	0.000
	(2)	0.215	0.000	0.359	0.223
	(3)	0.000	1.000	0.014	0.003
	(4)	0.610	0.000	0.430	0.492
	(5)	0.147	0.000	0.045	0.100

数値は各モデルの選択割合を表す

Z-test1; 5%水準, Z-test2; 10%水準のP値による選択

5 考察

本研究で新たに提案した G 推定のためのモデル選択基準 GQIC について、シミュレーションで評価を行ったところ、因果治療効果を規定する変数の変数選択において、有限標本での良い性能が示された。また、MEGA Study のデータ解析においてもシミュレーションとほぼ矛盾しない結果が得られた。

本研究で提案した手法の応用可能性について述べる。本研究では加法的な構造平均モデル（潜在結果変数の期待値の差をモデル化）における検討を行ったので、結果変数が連続量である場合や 2 値変数である場合のリスク差のスケールの検討において G 推定を用いた場合に幅広く適用可能である。G 推定法はノンコンプライアンスの補正だけではなく、観察研究において治療変数が時間と共に変化する場合の時間依存性治療の因果効果の検討、治療の直接効果・間接効果の検討にも用いられている。これらの問題にも提案した手法は直接的に応用可能であり、汎用性は高い。

本研究の限界については、時点以外の時間依存性共変量やそれまでの治療履歴と、その時点での治療変数との交互作用項を含んだモデルを検討できないことが挙げられる。しかしながら、その場合でも疑似尤度に直接的に基づく方法（尤度比検定、尤度比信頼区間、モデル選択基準）以外を用いた推測を行うことは可能である。

6 結論

本研究では、構造ネスト平均モデルにおける G 推定関数に対する疑似尤度関数を導き、その疑似尤度関数に基づいて種々の統計的推測を行うことを提案した。提案する方法によれば、ワルド、スコア、尤度比の 3 つのタイプの漸近検定や信頼区間の構成を行うことができる。また、AIC と類似したモデル選択基準 GQIC の導入を行った。シミュレーション研究の結果、提案するモデル選択手法は有限標本において高い性能を有することが示された。提案する方法を MEGA Study データに適用した結果、ノンコン

プライアンスの影響を補正したプラバスタチンのコレステロール抑制効果は投与してすぐに現れ、ほぼその値のまま持続することが示唆された。