

Asymptotic Relative Efficiency of Logistic and Poisson Models with both Response and Covariate Subject to Measurement Errors and its Application to Design of Epidemiological and Clinical Studies

長崎大学大学院生産科学研究科
三角 宗近

緒言

ロジスティックモデルは、遺伝子疾患関連研究において最もよく用いられ、ポアソン回帰は疫学追跡調査において相対リスクの推定及び有意性の検定のために広く用いられている。疫学研究では、測定値が正確に得られないことが多く、誤差を含む変数を使った統計的検定は検出力が下がることが知られている。研究計画段階で測定値に誤差を含むことを考慮して十分な検出力を得るように標本数を設定することが重要である。Pitmanの漸近相対効率 (asymptotic relative efficiency; 以下ARE) を用いることで、誤差を含む変数を用いて検定を行った場合の検出力の、正確な変数を用いた場合に対する相対的な損失がわかる。そのため、誤差の影響を考慮した場合、設定した検出力を得るためにどれぐらい多く標本を集めればよいかわかる。遺伝子疾患関連研究において、遺伝子型の特定は誤差を含み、ハプロタイプは直接決められずデータから推定されたものが用いられる。また、放射線影響研究所 (RERF) において1950年から行われている長期の原爆被爆者寿命調査では被曝による疾患の相対リスクを推定する際、被爆者の被曝放射線量推定値の誤差は常に考慮されている。一方、応答変数の誤差の影響に関心が集まることは少なかったが、癌臨床試験における薬の効果の判定は3次元の腫瘍縮小の有無を2次元画像で診断しており、誤診断の可能性は否定できない。また、途上国など、癌登録のシステムが未発達な地域では、報告されているデータよりもずっと多くの癌が発生している。本研究では、遺伝子疾患関連研究および疫学研究の計画をする際、応答変数と共変量の両方に誤差が存在する場合の検定への影響を評価し、適切な標本数を得るための方法確立した。さらに、実際の疫学研究を題材として、異なる測定誤差モデルにおいて提案手法が有効であることを示した。

以下、各章の概要を示す。

第1章では、疫学研究における測定誤差への対処、遺伝子癌臨床試験の実態、過少報告データの実態を紹介し、本研究の目的を明確にした。

第2章では、Pitman AREおよび測定誤差モデルの簡潔な概要を示した。

第3章では、ロジスティックモデルおよびポアソンモデルのどちらにおいても応答変数と共変量の両方で測定誤差を伴う場合のPitman AREが応答変数、共変量それぞれにおける正確な変数と誤差を含む観測された変数との相関係数の2乗の積で表されることを用いて、

正確な変数を用いる際の標本数にPitman AREの逆数を乗じることで測定誤差を考慮した必要な標本数を得ることを提案した。また、本研究で用いた測定誤差モデルごとにPitman AREを求めた。

第4章では、ロジスティックモデルにおいてPitman AREによって小標本でも提案手法によって適切な標本数が得られるかをシミュレーションによって検証した。シミュレーションでは一様分布からの乱数を共変量とし、様々な標本数、バックグラウンド奏効率、誤差率を仮定した。また、遺伝子癌臨床試験において測定誤差を考慮して必要な標本数を得るためのプログラムを設計した。

第5章では、ポアソンモデルにおいてPitman AREを求め、提案手法の妥当性を第4章と同様に検証した。さらに、RERF疫学研究における一般的な設定でPitman AREにより標本数を補正することで測定誤差の存在下でも検出力計算が適切に行われることを示した。

第6章では、第4章および第5章で行ったシミュレーションと応用例における結果を示した。

最後に、第7章において、本研究を総括するとともに、本研究で得られた結果を考察した。まず、ロジスティックモデルおよびポアソンモデルの両方において、Pitman AREを用いて得た標本数が適当であることが示され、実際の疫学研究でよく見られる設定においても適用可能であることが示された。応答変数、共変量の両方に誤差が存在する場合のPitman AREを求めた報告は今までにない。特に、ポアソンモデルにおいては未検出の症例の影響についてPitman AREを求めたのは初めてである。現実問題では、癌患者の症例報告がミスによってなされないことはあっても癌でない患者を癌と診断することは稀である。そのため、本研究で仮定した過少報告のデータ解析のためのポアソンモデルは疫学に限らず心理学や社会学など広く応用されている。このモデルの下でPitman AREを用いれば測定誤差の検定への影響をシンプルにそして適切に評価できることを示した。また、遺伝子疾患関連研究の計画の際に様々な大きさの誤差を想定し、広く利用できるプログラムを開発した。応答変数の誤識別についてdifferentialな誤差を仮定してAREを計算した結果、誤識別率 ϵ_0 、 ϵ_1 の検出力への影響が、バックグラウンド奏効率の値によって異なることが確認され、differentialな誤識別を考慮することで疫学研究に限らず、臨床検査などの診断検査器具の状況に即した研究計画に柔軟に対応することを可能にした。また、ポアソン回帰モデルが適用されるRERFの疫学研究において実際に用いられている加法測定誤差モデルでも正確な標本数が得られ、応用可能であることが示された。