

経験ベイズ縮約推定による 地点別事故危険性の評価 ～交差点における小学生事故を対象として～

松尾 幸二郎¹・宮崎 耕輔²・杉木 直³

¹ 正会員 豊橋技術科学大学 建築・都市システム学系 (〒441-8580 豊橋市天伯町雲雀ヶ丘1-1)
E-mail: k-matsuo@ace.tut.ac.jp

² 正会員 香川高等専門学校 建設環境工学科 (〒761-8058 高松市勅使町355)
E-mail: miyazaki@t.kagawa-nct.ac.jp

³ 正会員 豊橋技術科学大学 建築・都市システム学系 (〒441-8580 豊橋市天伯町雲雀ヶ丘1-1)
E-mail: sugiki@ace.tut.ac.jp

地点別の交通事故危険性評価においては、統計モデル分析により様々な説明変数が事故頻度へ与える影響を明らかにするアプローチがとられることが多い。しかし、地点の事故リスクには地点固有の非常に多様な要因が影響していることが考えられ、それらを全て説明変数として組み込むことは難しく、説明変数以外の要素は確率分布として考慮されることになる。従って、地点別の事故危険性を評価する上では、説明変数による平均的な影響と地点固有の特性による影響とのバランスを適切に考慮する必要がある。本研究では、経験ベイズ縮約推定によりこれを行うため手法を整理した。そして、ケーススタディとして、本手法を、生活道路交差点における小学生の事故危険性評価に適用した。その結果、過去5年間の交差点別の事故データを用いて将来5年間の事故頻度を予測する場合に、事故データのみによる予測および負の二項回帰モデルのみによる予測に比べ、経験ベイズ縮約推定による予測精度が若干向上した。また10地点より多くの危険地点を抽出する場合の効率性が向上した。

Key Words : spots' accident risks, empirical Bayese, statistical model, elementary students, traffic safety

1. はじめに

WHO¹⁾によれば、運動不足は健康における主要なリスク要素の1つとして認められている。特に子供においては、日常的な移動を徒歩や自転車で行うことが十分な運動を確保するために重要であり、それには子供による自由な移動可能性の程度を意味するChildren Independent Mobility (CIM)が関わっているとされる²⁾。わが国では大半の子供が、通学を中心に子供のみにより徒歩や自転車による移動を行っており、CIMも世界的に見て高い水準にある。しかし、子供の徒歩や自転車による移動には交通事故のリスクも伴うため、今後もCIMを高い水準に保つためには、適切な交通安全マネジメントを通して、子供の徒歩や自転車による移動の安全性を十分に確保する必要がある。

交通安全マネジメントにおいて、特に対策地点の抽出や対策の効果検証を行う上では、各地点の交通事故危険

性の評価を的確に行うことが重要である。地点別の事故危険性評価に関する既往研究では、統計モデル分析により様々な説明変数が事故件数へ与える影響を明らかにするアプローチがとられることが多い³⁾。しかし、地点の事故危険性には地点固有の非常に多様な要因が影響していることが考えられ、それらを全て説明変数として組み込むことは難しく、説明変数以外の要素は確率分布として考慮されることになる。従って、地点別の事故危険性を評価する上では、説明変数による平均的な影響と地点固有の特性による影響とのバランスを適切に考慮する必要がある。Hauer⁴⁾は、交通安全対策の事前事後評価の文脈で、経験ベイズ縮約推定によりこれを行うため手法を提案したが、地点別の事故危険性を適切に評価して対策地点の抽出という目的においても適用可能であると考えられる。

そこで本研究では、まず統計モデルおよび経験ベイズ縮約推定により地点別の事故危険性評価を行う手法につ

いて整理する。そして、愛知県豊橋市内の生活道路交差点における小学生の歩行中・自転車情報中の事故危険性を対象として、本手法を適用しその有用性を検証することも目的とする。

地点別の交通事故危険性の統計モデル分析には、事故リスク暴露量が重要とされる。すなわち本研究においては、各地点における児童の交通量や活動量である。しかしながら、これを広範囲に適切に計測することは困難である。そこで本研究では、豊橋市において整備している指定通学路データを用いることでその代替を試みる。

2. 研究方法

(1) 統計モデル

本研究では、地点別の小学生の歩行中および自転車乗用中における事故件数の期待値を「地点別の事故危険性」と定義する。従って、統計モデルとして計数モデルを用いる。シンプルな計数モデルとしてポアソン回帰モデルがあるが、誤差分布であるポアソン分布には期待値と分散が等しいという強い制約があるため、近年の事故件数モデリングにおいては、ポアソン回帰モデルの拡張の1つである負の二項回帰モデルを適用することが一般的となっている³⁾。本研究においても、負の二項回帰モデルを用いるが、経験ベイズ推定を用いて地点固有の要因を考慮した事故リスク評価を行うための手法について整理するため、以下ではポアソン回帰モデルにおける期待値がガンマ分布に従い確率的に変動するという特性を仮定することで負の二項回帰モデルに拡張されることを説明する。その上で、モデルに考慮した説明変数による平均的な影響（固定効果）では説明できない、地点固有の影響を考慮した地点別の事故危険性を、経験ベイズ法により評価する手法を示す。

a) ポアソン回帰モデル

ある対象期間中の地点*i*における事故発生件数 Y_i が、説明変数ベクトル \mathbf{x}_i および対応するパラメータベクトル \mathbf{b} の関数である期待値 λ_i を持つポアソン分布に従うとすると、

$$Y_i \sim \text{Poisson}(\lambda_i) = \text{Poisson}(\mathbf{x}_i, \mathbf{b}) \quad (1)$$

であり、地点*i*において対象期間中に y_i 件の事故が発生する確率は、

$$P_{\text{Poi}}(Y_i = y_i | \lambda_i) = \frac{\lambda_i^{y_i}}{y_i!} \exp(-\lambda_i) \quad (2)$$

となる。ここで、 Y_i の期待値と分散は、

$$E(Y_i | \lambda_i) = V(Y_i | \lambda_i) = \lambda_i \quad (3)$$

である。ただし、

$$\lambda_i = \exp(\mathbf{x}_i \mathbf{b}) = \exp\left(b_0 + \sum_k b_k x_{ik}\right) \quad (4)$$

とする。

b) 地点別の固有要因を考慮した事故危険性評価指標

上述のポアソン回帰モデルの期待値が、説明変数とパラメータの関数だけでは完全には決まらず、地点*i*固有の要因にも影響されると考え、期待値 λ_i の代わりとして、*shape* パラメータ $\alpha = \phi$ および *rate* パラメータ $\beta = \phi \lambda_i^{-1}$ を持つガンマ分布に従う確率変数 λ'_i を仮定すると、

$$\lambda'_i \sim \text{Gamma}(\alpha = \phi, \beta = \phi \lambda_i^{-1}) \quad (5)$$

となる。ここで、 λ'_i がある値 r_i になる確率は、

$$\begin{aligned} P_{\text{Gam}}(\lambda'_i = r_i | \alpha = \phi, \beta = \phi \lambda_i^{-1}) \\ = \frac{\phi^\phi \lambda_i^{-\phi}}{\Gamma(\phi)} r_i^{\phi-1} \cdot \exp(-\phi \lambda_i^{-1} r_i) \end{aligned} \quad (6)$$

であり、 λ'_i の期待値と分散は、

$$E(\lambda'_i) = \frac{\phi}{\phi \lambda_i^{-1}} = \lambda_i = \exp(\mathbf{x}_i \mathbf{b}) \quad (7)$$

$$V(\lambda'_i) = \frac{\phi}{\phi^2 \lambda_i^{-2}} = \frac{\lambda_i^2}{\phi} \quad (8)$$

となる。事故件数期待値 λ'_i がガンマ分布に従って決まり、事故発生件数 Y_i がその λ'_i をパラメータとするポアソン分布に従って発生するという構造は、階層モデルであると解釈ができる。そこで、地点*i*において実際に y_i 件の事故が発生したという情報の下で、 λ'_i がある値 r_i である確率をベイズの定理により求めると、

$$\begin{aligned} P(\lambda'_i = r_i | Y_i = y_i; \lambda_i, \phi) \\ = \frac{P_{\text{Poi}}(Y_i = y_i | \lambda_i) P_{\text{Gam}}(\lambda'_i = r_i | \phi, \lambda_i^{-1})}{P(Y_i = y_i | \lambda_i, \phi)} \end{aligned} \quad (9)$$

となる。ここで、 λ_i および ϕ は超事前パラメータである。さらに、ガンマ分布はポアソン分布の自然共役分布であり、事後分布もガンマ分布となることが知られており、

$$\begin{aligned} P(\lambda'_i = r_i | Y_i = y_i; \lambda_i, \phi) \\ = P_{\text{Gam}}(\lambda'_i = r_i | \phi + y_i, \phi \lambda_i^{-1} + 1) \end{aligned} \quad (10)$$

となることが示される。また、その期待値は、

$$\begin{aligned} E(\lambda'_i | \phi + y_i, \phi \lambda_i^{-1} + 1) \\ = \frac{\phi + y_i}{\phi \lambda_i^{-1} + 1} \end{aligned} \quad (11)$$

となる。この λ'_i の期待値は、地点固有の要因も考慮した事故危険性評価指標の1つとして用いることができると考えられる。

c) 負の二項回帰モデルを通じた経験ベイズ推定

式(11)の事故危険性評価指標は、超事前パラメータ λ_i および ϕ に依存する。従って、本研究ではこれら超事前パラメータを経験ベイズ法により推定する。すなわち、周辺尤度である、

$$\begin{aligned} P(Y_i = y_i | \lambda_i, \phi) \\ = \int_{-\infty}^{\infty} P_{\text{Poi}}(y_i | r_i) P_{\text{Gam}}(r_i | \phi, \lambda_i^{-1}) dr_i \end{aligned} \quad (12)$$

を最大化する λ_i および ϕ を用いる。上式をさらに展開し、

$p_i = \phi / (\phi + \lambda_i)$ とおくと,

$$P(Y_i = y_i | \lambda_i, \phi) = \binom{y_i + \phi - 1}{y_i} (p_i)^{\phi} (1 - p_i)^{y_i} \quad (13)$$

となり, Y_i がパラメータ λ_i および ϕ を持つ負の二項分布に従うことが示される. すなわち,

$$Y_i \sim NB(\lambda_i, \phi) = NB(\mathbf{x}_i, \mathbf{b}, \phi) \quad (14)$$

であり, Y_i の期待値と分散は,

$$E(Y_i | \lambda_i, \phi) = \frac{1 - p_i}{p_i} \phi = \lambda_i = \exp(\mathbf{x}_i, \mathbf{b}) \quad (15)$$

$$V(Y_i | \lambda_i, \phi) = \frac{1 - p_i}{p_i^2} \phi = \lambda_i + \frac{\lambda_i^2}{\phi} \quad (16)$$

となる. ここで, ϕ は分散パラメータ (dispersion parameter) と呼ばれ, この値が大きいほど Y_i の分散 (残差逸脱度) が小さいことを示しており, $\phi = \infty$ の時には分散が期待値と等しくなり, すなわち負の二項分布がポアソン分布と一致することが分かる.

以上より, 地点別の実際の事故件数 y_i と説明変数ベクトル \mathbf{x}_i を用いて負の二項回帰モデルを推定した上で, モデルにより得られる分散パラメータ ϕ^{NB} および地点別の事故件数期待値 λ_i^{NB} を超事前パラメータとして用いることで, λ_i' の経験ベイズ推定が可能となる. λ_i' の経験ベイズ推定量を λ_i^{EB} とすると,

$$\lambda_i^{EB} = \frac{\phi^{NB} + y_i}{\phi^{NB} \lambda_i^{NB-1} + 1} = \left(\frac{1}{1 + \frac{\lambda_i^{NB}}{\phi^{NB}}} \right) \lambda_i^{NB} + \left(\frac{1}{1 + \frac{\lambda_i^{NB}}{\phi^{NB}}} \right) y_i \quad (17)$$

となることが示される. これより, この経験ベイズ推定量は, 負の二項回帰モデルにおける説明変数の平均的な影響により算出された地点 i の事故件数期待値 λ_i^{NB} と実際に発生した事故件数 y_i との重み付き平均であることが分かる. そして, 負の二項回帰モデルの分散パラメータである ϕ^{NB} が大きいほど, すなわち負の二項回帰モデルによる残差逸脱度が小さいほど, λ_i^{NB} に大きな重みがつくことが分かる. このような推定量は, 紺田⁵⁾によれば, 「標本での推定量を全体の推定量に縮約しており, 個別の推定において全体の力による説得力の借用をしている (Borrowing strength from the ensemble). このような推定量は縮約推定量と呼ばれる。」ものである.

(2) データ概要

a) 対象地域

対象地域は, 愛知県豊橋市全域とした. 豊橋市には小学校が52校あり, 2019年現在の小学生の数は約2万人である. 10名程度の自転車通学児童を除くほぼ全ての児童が徒歩で通学している.

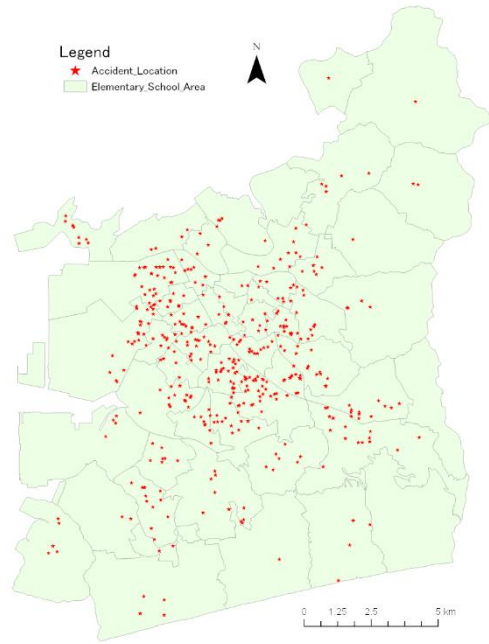


図-1 対象地域における小学生事故発生箇所

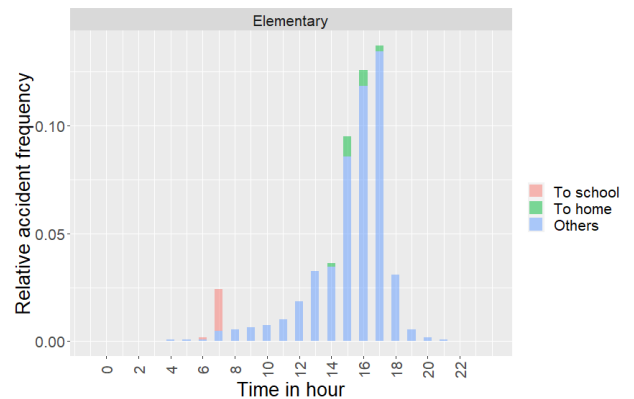


図-2 対象地域における小学生事故の目的別・時間帯別構成

b) 事故データ

本研究で用いた事故データは, 2009年~2018年までの10年間に豊橋市内で発生した小学生事故 (第1当事者もしくは第2当事者が小学生である人身事故, 1,072件) である. 図-1に小学生事故の発生箇所分布を, 図-2に小学生事故の目的別・時間帯別構成を示す. 夕方の通学目的以外での移動中の事故が多いことが分かる. このうち, 各生活道路交差点中心から半径50m以内で発生した事故341件 (小学生事故全体の31.8%) を交差点事故とした. ここで, 生活道路交差点は, Zmap-Area II (ゼンリン社) において道路種別が「高速自動車国道, 一般国道, 一般都道府県道, 主要地方道, 主要一般道」以外の道路のみによる3肢以上の交差点として定義した. 対象となる交差点8,361交差点のうち, 対象期間10年間で, 277地点 (3.3%) で1件, 24地点 (0.29%) で2件, 4地点 (0.048%) で3件, 1地点 (0.012%) で4件の小学生事故が発生して

いる。なお、前半5年間の小学生事故186件を地点別事故危険性の推定に用い、後半5年間の155件を、推定結果の評価に用いた。

c) 小学生事故リスク暴露量データ

地点別の事故危険性の統計モデル分析には、事故リスク暴露量が重要とされる。すなわち本研究においては、各交叉点における児童の交通量や活動量である。しかしながら、これを広範囲に計測することは困難である。一方、豊橋市教育委員会では、市内全小学校の通学団ごとの指定通学路および集合箇所と児童数を指定通学路データとして地理情報化している。そこで本研究では、この指定通学路データ（2018年度版）に基づき、交叉点別事故リスク暴露量の代理変数として以下の指標を算出した。なお、指定通学路データは近年の年度による大きな違いはない。

- **周辺通学路児童数**：各対象交叉点中心から半径15m以内、50m以内、100m以内を通過する指定通学路の児童数（図-3参照）
- **周辺集合箇所児童数**：各対象交叉点中心から半径15m以内、50m以内、100m以内にある集合箇所の児童数（図-3参照）
- **最近接通学路距離**：各対象交叉点中心から最も近い指定通学路までの距離（図-4参照）

d) その他のデータ

小学生事故件数には自動車交通量の多寡による影響も大きいと考えられる。そこで、各対象交叉点における自動車交通量の多寡を考慮するため、本研究ではパイオニア社の一般車両プローブデータ2016年1年分を用いた。対象交叉点別に、交叉点に流入する方向のプローブ車両通過量を全肢分総和したものを、各対象交叉点のプローブ車両通過量とし（図-5参照）、自動車交通量の多寡を表現する指標とした。

また、道路交通環境条件に関するその他の説明変数として、対象交叉点の交叉点肢数（道路接続本数）、信号の有無、国土数値情報の1/10細分メッシュ区画（100mメッシュ）別の都市地域土地利用種別を用いた。

3. 結果

(1) 負の二項回帰モデル推定結果

交叉点別の小学生事故発生件数を目的変数とした負の二項回帰モデルの推定結果を表-1に示す。

小学生の事故リスク暴露量の代理変数としては、各対象交叉点中心から半径15m以内を通過する指定通学路の児童数を用いたモデルが最も適合度が高くなった。これは、小学生の事故が通学路上あるいは通学路付近で発生しやすいことを示している。上述したように、小学生事故は通学目的以外での移動中によるものが大半であるこ

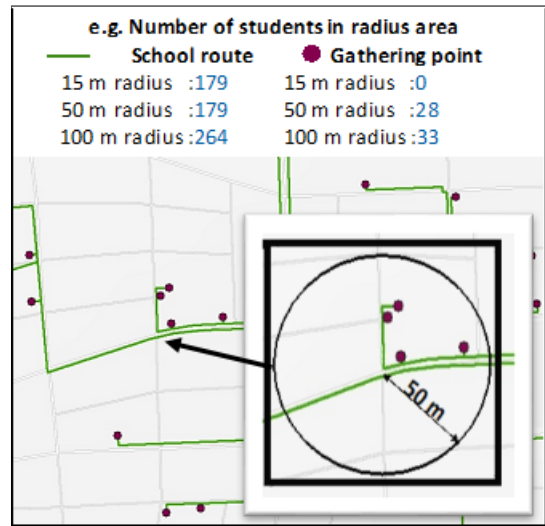


図-3 交叉点付近通学路児童数・集合箇所児童数の算出例

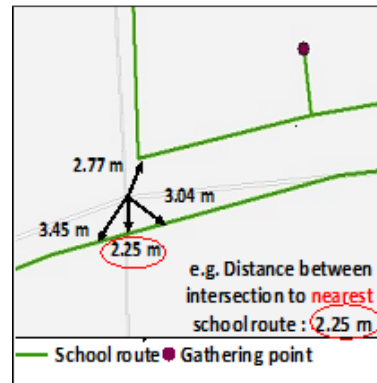


図-4 最近接通学路距離の算出例

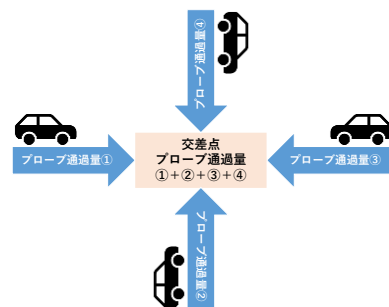


図-5 交叉点別プローブ交通量の集計方法

とを踏まえると、小学生は通学目的以外での移動にも普段から慣れている通学路を利用する傾向があり、その結果として通学路付近の事故件数が高くなっている可能性が示唆される。これはより詳細な研究が必要であるが、もしそうであれば、通学路に重点的に交通安全対策を実施していくことは理にかなっていると考えられる。また、指定通学路データが、小学生の事故リスク暴露指標の代理変数としても有用であることを示唆している。

それ以外の説明変数については、プローブ車両通過量が多いほど小学生事故危険性が高くなる傾向にあること、

3枝交差点に比べ4枝交差点の方が事故危険性が高い傾向にあること、居住用建物土地利用内の交差点において事故危険性が高まる傾向にあることが示された。

(2) 経験ベイズ縮約推定結果

図-6は、前期5年間の事故件数別に、事故危険性推定値（5年間の期待事故件数）について、横軸に負の二項回帰モデル推定値、縦軸に経験ベイズ推定値をとってプロットしたものである。45度線もあわせて引いている。経験ベイズ推定により、前期事故件数が0件の地点については期待事故件数が低減し、前期事故件数が1件以上の地点については期待事故件数が増大していることが確認できる。また、同じ前期事故件数であっても負の二項回帰モデルによる推定値が高いほど、その変化の大きさは大きいことが確認できる。

経験ベイズ縮約推定の有用性評価のため、前期5年の事故件数、負の二項回帰モデル推定値、経験ベイズ推定値それぞれに対して、後期5年の事故件数との相関係数を算出した結果を、表-2に示す。そもそもの小学生事故件数が多くないこともあり、相関係数自体は全体として高くはないが、前期事故件数との相関係数に比べ、経験ベイズ推定により相関係数が若干向上していることが分かる。

さらに、事故危険地点の効率的な抽出という側面から、経験ベイズ推定値の有用性を検証するため、以下の手順で、効率性指標を算出した。ここでの効率性の考え方は、「抽出した事故危険地点1地点当たりの後期事故発生件数が多いほど効率性が高い」というものである。

- ① 前期事故件数、負の二項回帰モデルによる事故危険性推定値、経験ベイズ推定の事故危険性推定値のそれぞれに基づき、各地点を危険度の大きい順に順位付け
- ② ①で付与した3基準の順位に基づき、各順位までの地点の後期事故件数の累積和を算出
- ③ ②で算出した3基準の後期事故件数の累積和を各順位までの地点の地点数で割ったものを効率性とする

図は、前期事故件数基準、負の二項回帰モデル推定値基準、経験ベイズ推定値基準の効率性を示したものである。危険性上位10地点までの抽出においては、前期事故件数基準による方が効率が良いが、それ以上の事故危険地点を抽出する場合には、経験ベイズ推定値基準の方が効率が良いことが分かる。

4. おわりに

本研究では、経験ベイズ縮約推定により地点別の事故危険性評価を行う手法について整理した上で、愛知県豊

表-1 小学生事故数モデルの推定結果（負の二項回帰）

説明変数	パラメータ	p値
定数項	-5.00	<0.001
周辺通学路児童数（15m）	-0.00568	<0.001
プローブ車両通過量	0.000106	0.019
3枝交差点（基準）	0	-
4枝交差点	0.744	<0.001
5枝以上交差点	0.751	0.48
高層建物土地利用	1.32	0.028
低層建物土地利用	0.845	<0.001
低層建物（密集）土地利用	0.721	0.028
上記以外土地利用（基準）	0	-
分散パラメータ ϕ		0.343
サンプルサイズ（対象交差点数）		8361
McFadden's ρ^2		0.0705
AIC		1717

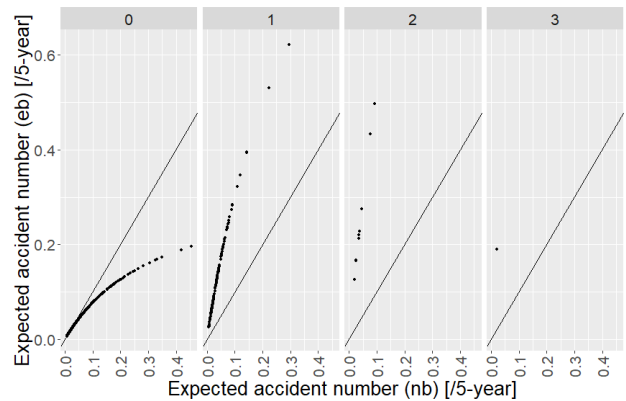


図-6 前期事故件数別（左から0~3）の事故危険性推定値比較（横軸：負の二項回帰モデル推定、縦軸：経験ベイズ推定）

表-2 事故危険性推定値と後期事故件数との相関係数

対・前期事故件数	0.131
対・負の二項回帰推定値	0.084
対・経験ベイズ推定値	0.144

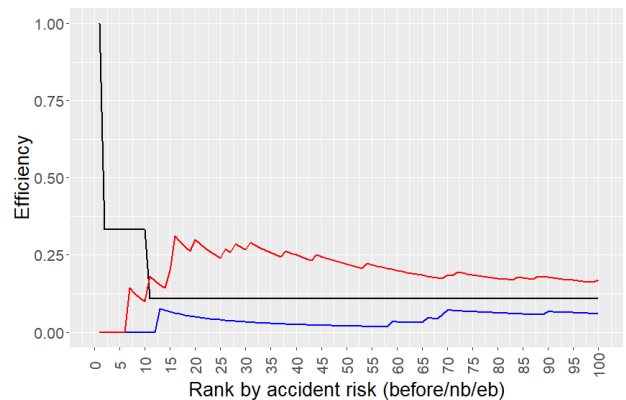


図-7 危険地点抽出の効率性比較（黒：前期事故基準、青：負の二項回帰モデル推定値基準、赤：経験ベイズ推定値基準）

橋市内の生活道路交差点における小学生事故危険性を対象として、本手法を適用しその有用性の検証を行った。結果をいかにまとめる。

- 負の二項回帰モデル分析の結果、小学生の事故が通学路上あるいは通学路付近で発生しやすいことが示され、小学生は通学目的以外での移動にも普段から慣れている通学路を利用する傾向があり、その結果として通学路付近の事故件数が高くなっている可能性が示唆された。
- 経験ベイズ縮約推定による各地点の小学生事故危険性の評価により、将来の事故発生件数の予測力が、過去の事故発生件数基準と比較して、若干向上した。
- さらに、10地点より多い危険地点を抽出する場合は、経験ベイズ推定値基準による方が効率性が良くなることが示された。

以上のことから、経験ベイズ縮約推定による地点別事故危険性評価の一定の有用性が示されたと考えられる。経験ベイズ縮約推定の考え方は、実際に発生した値とモデルを通した説明変数による平均的な影響との混合により予測力を向上させることである。従って、負の二項回帰モデルを改良して説明力を向上させることで、経験ベイズ推定による有用性もさらに向上すると考えられる。

謝辞：本研究は JSPS 科研費20K04743の助成を受けて実施した研究の成果である。ここに記して謝意を表す。

参考文献

- 1) WHO: Global Recommendation on Physical Activity for Health, WHO, Geneva, 2010
- 2) Cooper, A. R., Andersen, L. B., Wedderkopp, N., Page, A. S. and Froberg, K.: Physical activity levels of children who walk, cycle, or are driven to school. American Journal of Preventive Medicine, vol.29(3), pp.179-184, 2005
- 3) Lord, D. and Mannering, F.: The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives, Transportation Research Part A: Policy and Practice, vol.44(5), pp.291-305, 2010
- 4) Hauer, E.: Observational Before-after Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety, Emerald Group Pub, Oxford, 1997
- 5) 紺田広明：自殺における相対リスクのベイズ推定—経験ベイズ法と階層ベイズ法による縮約推定量の構成—, 社会的信頼学, vol.2, pp.1-34, 2014

(2020. 10. 2 受付)

AN EMPIRICAL BAYESIAN SHRINKAGE ESTIMATION TO EVALUATE ACCIDENT RISKS BY LOCATIONS: AS A CASE STUDY FOR ELEMENTARY SCHOOL STUDENTS ACCIDENTS AT INTERSECTIONS

Kojiro MATSUO, Kosuke MIYAZAKI and Nao SUGIKI