

一般化線形モデルを用いた幹線道路における自転車事故の事故リスク分析

都市基盤計画分野 亀井省吾

Abstract

自転車事故分析に関して、道路リンクを単位とした走行空間の事故リスクの違いをリスクモデルにより評価する研究事例は海外では報告されているものの、日本では事例が少ないのが現状である。また、事故件数のように、被説明変数が離散型であることを踏まえたポアソン回帰モデルによる要因分析では、観測した分散が推定した分散より大きくなる過剰分散問題があり、実事故データから有益な情報を抽出する方法論の課題もある。そこで、本研究では、この課題に対応できる一般化線形モデルを用いて、その特性と自転車事故に影響を与える要因の把握を目的として、分析を行った。本研究の成果として、一般化線形モデルの有用性が確認された。また、事故件数に影響を与えるものは、交通量と交差点であり、自転車通行施設の違いは影響を与えないことが明らかになった。

1. はじめに

近年、健康志向の高まりなどによる自転車使用の促進に伴い、自転車が安全・快適に走行することを目的として、自転車レーンや自転車通行帯などの自転車走行空間の整備が進んでいる。これらの空間が実際に安全かどうかを判断するためには、客観的な事故分析結果に基づいた問題点の把握と対策立案が不可欠である。

自転車事故に関しては、事故統計データを用いた分析事例や現地観測に基づいた報告がある¹⁾。特に交通量の変化に伴う影響や、自転車走行空間による事故率の違いについては、海外において分析事例²⁾が見られる。一方、国内では、利用可能な統計データも少なく、事故が多い特定箇所における分析事例はみられるものの、自転車走行空間の安全性の比較についてほとんど行われていないのが現状である。

また、道路空間の客観的評価手法の1つとして、道路リンクを単位とした事故分析モデルの構築があげられる。特設箇所(点)よりも道路リンク(線)を単位とすることで、マクロ的な分析が可能と考えられる。これまでにも、同様の手法で、平均=分散と仮定するポアソン回帰モデルを用いた事故統計分析が行われてきたが、用いるデータが原因で生じる過剰分散問題(平均<分散)により、パラメータ推定にバイアスがかかるなど、実際の事故データから有益な情報を抽出する方法論にも課題が生じている²⁾。

以上より、国内では、道路リンクを対象とした自転車事故分析はほとんど行われていないので、自転車走行空間の違いを評価できる道路リンクを対象としたマ

クロ的観点の分析が必要である。また、自転車事故をモデル式を用いて分析する際には、過剰分散の影響を緩和できるモデル式を用いる必要がある。

そこで、本研究では、以下の2つを本研究の目的として、兵庫県下の幹線道路で生じた自転車事故を対象に、過剰分散問題に対応するために適用分布が変更可能である「一般化線形モデル」を用いて、事故リスク分析を行った。

- ①実際の日本のデータ(道路空間データ、事故データ)を一般化線形モデルにあてはめ、その有用性と本研究で用いたデータと過剰分散の関係を明らかにする。
- ②道路リンク毎に、自転車事故の潜在的危険に影響を与える要因を明らかにする。

2. 研究対象と分析手法

2.1 本研究の研究対象

本研究では、兵庫県下の幹線道路で生じた自転車事故を分析対象とする。そこで、H17 道路交通センサスと H17 兵庫県事故データを使用し、モデル構築に必要な情報を得た。

本研究で取り扱ったデータは、H17 道路交通センサスに記録のある兵庫県下の幹線道路のうち、一般国道と主要地方道に関しては全てを網羅した。しかし、本研究は事故対象を自転車関連事故とするため、H17 道路交通センサスにおいて、主に自動車専用道路からなる自転車交通量が観測されなかった 94 区間については分析の対象外とした。残りの 612 区間を分析の対象とし、それをまとめたものを以下に示す(表 1)。

自転車通行施設については、兵庫県内に、自転車通行帯はほとんどないことがわかる。そこで、本研究では「自転車通行帯」は分析の対象には考慮せずに、「自転車歩行車道」「歩道あり道路」「歩道無し道路」の3種類を分析の対象とすることにした。

また、事故データに関しては、H17 兵庫県事故データから、同区間内の自転車関連事故を含む主要な事故類型について GIS を用いて集計した (表 2)。

モデル式に代入する説明変数は、道路交通センサスの調査項目の中から、暴露量、通行施設、錯綜条件、走行条件に関わる項目を選定した (表 3)。

2.2 本研究の研究方法

(1) 事故リスクの定義

一般に、事故が生じるのには様々な原因があり、ハ

表 1 分析対象路線データ

	リンク数	区間延長	自転車通行帯延長	自転車歩行者道延長	歩道あり道路延長	歩道なし道路延長
一般国道	267	1182	0	586.5	160.2	430
主要地方道	345	1647.8	0.1	521.7	191.3	934.7
計	612	2829.8	0.1	1108.2	351.5	1364.7
幹線道路	1565	6292.8	1.7	1724.2	581.5	3985.4
県内道路 総延長に 占める割合	39%	45%	6%	64%	60%	34%
自転車通行 施設構成率			0%	27%	9%	63%

(単位: km)

表 2 対象事故件数データ

事故類型	一般国道	主要地方道	計	
自転車×車両	軽傷	1349	823	2172
	重傷	145	80	225
	死亡	9	11	20
	計	1503	914	2417
自転車×自転車	110	68	178	
自転車×歩行者	15	10	25	
車両×車両	6514	4329	10843	
車両×歩行者	469	473	942	
車両単独	432	401	833	
二輪×車両	1391	763	2154	

表 3 選定した説明変数データ

暴露量	空間条件		
	通行施設	錯綜条件	走行条件
・平日自動車交通量 ・平日自転車交通量 ・平日自動二輪交通量 ・平日歩行者交通量 ・区間延長	・歩道なし ・歩道あり ・自転車歩行者道	・信号有無別 交差点数	・路肩幅員 ・歩道代表幅員 ・車道幅員 ・指定最高速度

ードウェア、人間、システムがあげられる³⁾。これらを自転車事故に当てはめると、事故の原因は「交通具」、「運転手」、「走行環境」にわけることができる。交通事故統計原票マニュアルでも「車両的要因」「人的要因」「環境的要因」の3つに分類されており、その細分類も含めたものが以下ようになる (表 4)。

事故が生じるには、暴露量で説明できない潜在的な危険がある。そこで、「事故リスク」は、この潜在的な危険 (事故) を発生させる確率を意味し、事故が事故当事者に与える被害の程度も含むものと本研究では定義する。

表 4 事故リスクに作用する要因

原因	ハードウェア	人間	システム
	交通具	運転手	走行環境
原票マニュアルにある分類	車両的要因 ・整備不良 ・状態不良	人的要因 ・発見の遅れ ・判断の誤り ・操作上の誤り	環境的要因 ・道路的要因 ・交通環境的要因
具体例	・自転車故障 ・自動車故障 etc	・運転ミス ・信号無視 ・不注意 etc	・交差点数 ・自転車通行施設 ・幅員 etc

(2) モデルの概要

本研究では、空間の持つ潜在的な危険度 (事故リスク) と暴露量 (試行回数=交通量) とにわけ、その積によって事故が顕在化すると考える。言い換えれば、事故件数、事故リスク、交通量を独立した変数として取り扱い、観測された事故件数および交通量から未知の事故リスクを推定するものである。また交通量との関係が線形とは限らないので、式 1 のような乗数モデルを用いる。この式中の α は、暴露量とは独立した事故リスクを表す未知のパラメータで、潜在的な事故の危険度を表す。 β は事故件数と交通量との線形、非線形関係を表し、 $\beta=1$ 、 $\beta>1$ 、 $\beta<1$ の時にそれぞれ線形、指数、対数関係となる。

$$Y_i = \alpha \times Q_i^\beta \dots (\text{式 1}) \quad Y_i \sim Po(\mu_i)$$

この両辺を対数にとり、事故リスクの変化を表す説明変数を加えたものが式 2 となる。

$$\log \mu_i = \log \alpha + \beta \log Q_i + \sum \beta X \dots (\text{式 2})$$

μ : 事故件数期待値 (件/年) Y : 実事故件数

α : 事故リスク Q : 交通量

X : 事故リスクの変化要因 i : リンク番号

交通事故件数は平均と分散が等しい離散型分布のポアソン分布に従うと仮定することが一般的であるが、実際に観測した平均および分散と推定した分散が等し

くなることはなく、観測分散>平均となる過剰分散という現象が生じることが確認されている。

そこで過剰分散問題をモデル上で緩和するために、被説明変数に適用する分布が変更可能な一般化線形モデルを使用し、負の二項分布（平均≠分散）を適用した分析を行う。

この分布を適用して分析を行う際には、過剰分散の程度を表すものとして式3に示す過剰分散パラメータΦを求めることができる。Φが大きいほど過剰分散し、過剰分散パラメータが0になると平均と分散が等しくなり、ポアソン分布に従うことになる。過剰分散パラメータがおよそ1.6を超えると統計的に有意となり、過剰分散が確認されたことになる。

$$\sigma^2 = \mu_i(1 + \phi\mu_i) \quad (\sigma : \text{標準偏差} \quad \mu : \text{平均}) \dots (\text{式3})$$

離散モデル全体の精度を表す指標として、対数尤度（log likelihood）と赤池情報基準（AIC）を用いた。対数尤度に説明変数の数の影響を考慮したAICの値が小さいほど、モデルの精度が向上したと判断できる。

$$\log \text{likelihood} = -n\mu + \log \mu \times \sum_{i=1}^n x_i - \sum_{i=1}^n \log x_i!$$

$$AIC = \frac{-2 \times (\log \text{likelihood}) + 2 \times (\text{変数の数})}{n}$$

3. データセットの基礎統計

ここで、H17道路交通センサスとH17兵庫県事故データより、基礎統計を示し、データセットから読み取れる自転車事故の傾向を把握する。ただし、ここでは交通量の影響は考慮していない。

H17兵庫県事故データより、交通事故発生箇所に着目して、各事故類型を比較したものを以下に示す（図1）。各事故類型の中でも、「自転車×車両事故」においては、交差点事故：76.4%と他の事故類型と比較すると非常に交差点での事故が生じやすく、単路部での事故が生じにくいことがわかる。

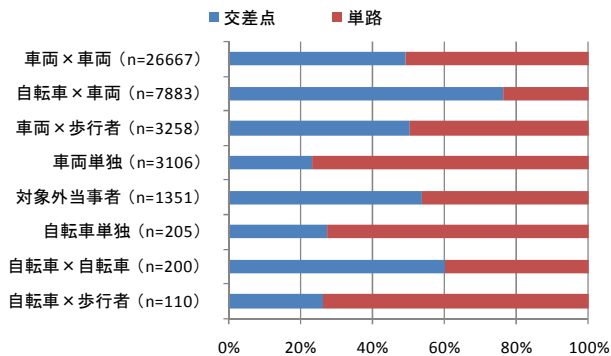


図1 事故類型別・事故発生箇所割合

次に、H17道路交通センサスとH17兵庫県事故デー

タのクロス統計より、各種道路における単位距離あたりの事故件数割合を事故類型別に比較したものを以下に示す（図2）。いずれの事故類型に関しても、歩道あり道路での事故割合が高いことがわかる。「自転車×車両」に注目すると、先ほどの結果より、交差点事故が多いため、単路部の車道と歩道を分離しても事故抑制に繋がりにくいと考えられる。

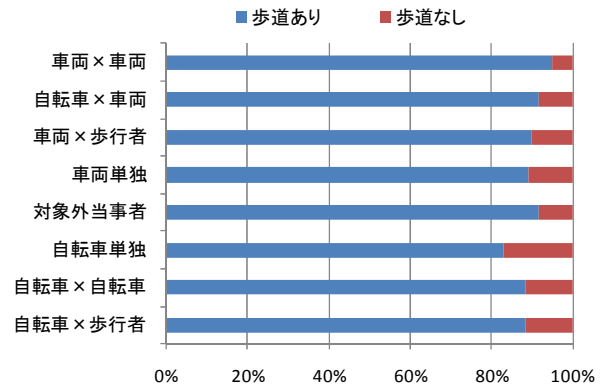


図2 事故類型別・歩道有無割合（単位距離あたり）

次に、交差点に注目し、交差点の信号有無まで考慮した基礎統計を示す。そこで、各種単位交差点あたりの事故件数を事故類型別に比較した（図3）。

結果より、「車両×車両」と「自転車×車両」の事故件数に注目する。「車両×車両」は信号交差点で2.0件、信号なし交差点で0.6件となり他の事故類型よりも大きく信号交差点の影響を受けることがわかる。「自転車×車両」は信号交差点で0.6件、信号なし交差点で0.34件と信号交差点の影響の方が大きい、「車両×車両」と比較するとその影響が小さいことがわかる。

これより、「車両×車両」は信号交差点の影響を受けやすく、「自転車×車両」は「車両×車両」よりも信号なし交差点の影響を受けやすいと考えられる。

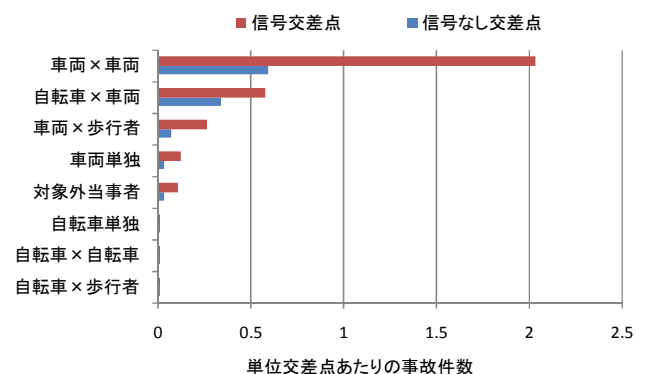


図3 事故類型別・各種単位交差点あたり事故件数

5において、交通量の影響を考慮した、自転車事故リスクモデルの分析結果から自転車事故に影響を与える要因を調べ、基礎統計から得られた結果と比較をする。

4. 対象データセット・モデルの検証

ここでは、まず「適用分布の比較」と「交通量と事故件数の関係」に関する分析を行い、一般化線形モデルの有用性に関する分析を行う。

そして、対象データの「サンプリング率」と「説明変数」の2つの側面から過剰分散パラメータの変化を検証し、過剰分散の原因を調べた。

4.1 過剰分散を考慮したモデルに関する分析

過剰分散に対応できるか検証するために、「負の二項分布」と「ポアソン分布」を本研究のデータセットで構築したモデル式に適用し、モデル全体の精度を AIC の値を用いて比較する。

自転車×車両事故を対象に、事故件数と交通量の関係を表すモデル式を暴露量に①（自転車交通量+車両交通量）×区間延長（1変数）、②（自転車交通量+車両交通量）、区間延長（2変数）、③自転車交通量、車両交通量、区間延長（3変数）という3パターンをとり、分析を行った。AICの結果を以下に示す。（図4）。

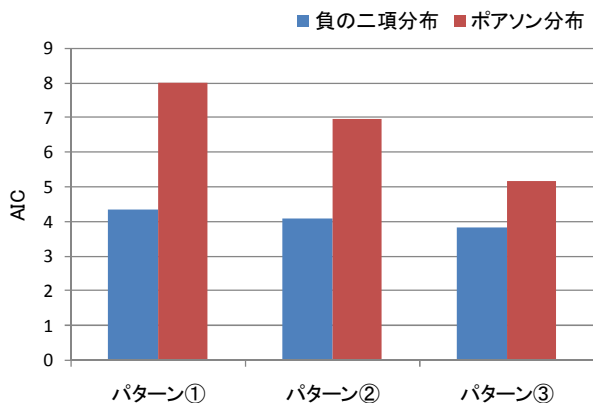


図4 各パターン AIC 比較結果

結果より、いずれのパターンにおいても、負の二項分布を用いた方が、分析の精度は向上した。これは説明変数の数を増やしていても同様の傾向を示すことから、負の二項分布を使用した方がより精度の良い分析が可能であることが明らかになった。そのため、これ以降の分析に関しては、負の二項分布を適用した分析のみを行っていく。

また、先ほどの3パターンにおいて、負の二項分布を用いた分析結果を以下に示す（表5）。結果より、すべてのパターンにおいて、交通量の次数は1より小さくなったため、事故件数と交通量の関係は非線形（対数関係）といえることができる。また、分析結果から得られた推定式を用いて、自転車交通量と予測事故件数の関係を図に示す（図5）。その際、実事故件数と線形モデルによる予測事故件数も記載した。

図5より、交通量と事故件数の関係を線形と仮定すると、予測事故件数に関して、交通量の少ない範囲では過少評価してしまい、交通量の多い範囲では過大評

価してしまう問題が見て取れる。

表5 負の二項モデルによる分析結果

	パターン①		パターン②		パターン③	
	係数	t 値	係数	t 値	係数	t 値
定数	-6.77	-24.2 **	-7.70	-27.3 **	-9.45	-13.1 **
(自転車交通量+車両交通量) × 区間延長	0.75	30.54 **				
自転車交通量+車両交通量			0.88	34.14 **		
区間延長			0.41	12.75 **	0.78	20.99 **
自転車交通量					0.73	34.23 **
車両交通量					0.57	14.75 **
log likelihood	-219.71		-2089.76		-893.79	
AIC	4.342		4.070		3.842	

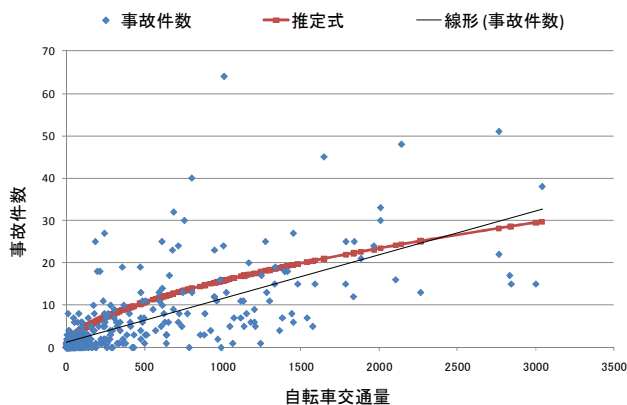


図5 自転車交通量と事故件数の関係

4.2 過剰分散に関する分析

(1) サンプリング率との関係

10~75%のサンプリング率7パターンごとにランダムにデータを10回抽出し、4.1のパターン③と同様のモデル式を用いて分析を行い、10回の過剰分散パラメータの平均と分散を比較した結果を以下に示す（図6）。

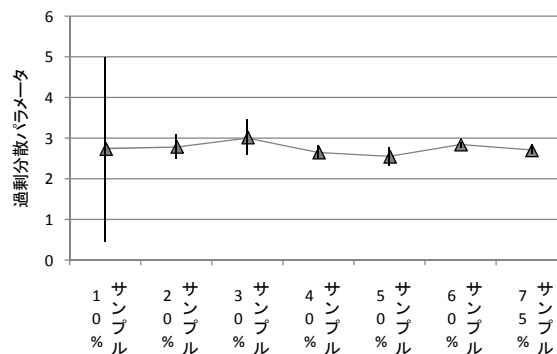


図6 サンプリング率別過剰分散パラメータ

結果より、いずれのパターンにおいても、過剰分散が生じていることが確認された。これより、本研究のデータセットはいかなるサンプリング率でも過剰分散

が生じていることが明らかになった。また、サンプリング率 10%の時だけ、過剰分散パラメータの分散が非常に大きな値となった。そこで、10%のデータセットを調べたところ、区間延長の影響を受けている可能性があると考えられた。

(2) 説明変数との関係

ここでは、説明変数の種類を増やすことにより、過剰分散パラメータの変化を検証をした。説明変数は、交差点密度、信号有無別交差点密度、歩道なし道路延長率を順に説明変数として、追加していく。また、その際、データセットのサンプルを区間延長が 4.3km 以上以下の 2 グループに分けて、それぞれ分析を行った。また、事故類型の違いが過剰分散に与える影響を確認するため、「車両×車両」についても、同様に行った。結果を以下に示す (図 7)。

結果より、事故類型の違いでは、自転車×車両のほうが過剰分散パラメータが大きく、加えて区間延長が長い場合にその傾向が強くなった。このことは、区間延長の長い場合に事故発生地点が不均質になるためと考えられ、特に自転車×車両事故は車両×車両事故よりも特定箇所集中するものと考えられる。

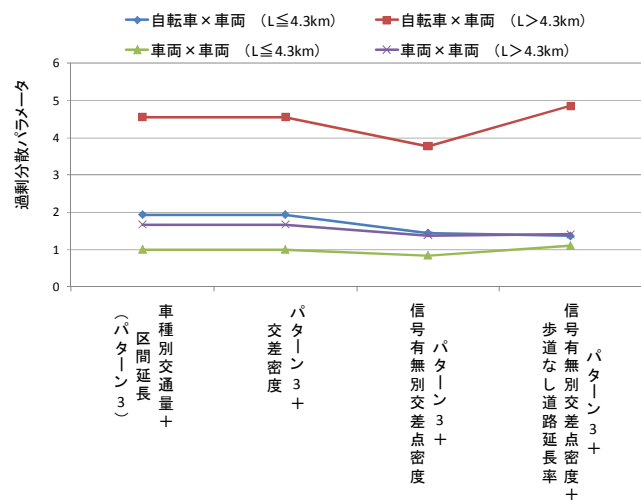


図 7 事故類型・区間延長別 過剰分散パラメータ

4.1 の結果より、負の二項分布を考慮でき、非線形モデルを表現できる一般化線形モデルの有用性が確認された。4.2 の結果から、本研究のデータセットはサンプリング率、説明変数の数にかかわらず、常に過剰分散が発生していることが確認された。その要因として、区間の分割方法や区間内での事故発生箇所の均質性が関与していることが示唆できた。しかし、本研究でこの問題を解決することが目的でないことと、分析に用いたセンサデータを使ってこの問題を解決することは困難なため、以降の分析においては、同じデータセットに対して、過剰分散問題を緩和できる負の二項分布を用いたモデルで分析を行うこととした。

5. 自転車事故のリスク分析

5.1 自転車通行施設の比較

ここではどのような自転車通行施設であれば安全なのかを調べることを目的とし、4.1 のパターン③にさらに通行施設に関する変数を追加し、分析を行った。

分析結果 (表 6) より、暴露量に関する変数である自転車交通量、車両交通量、区間延長は有意であった。しかし、自転車通行施設に関する変数である歩道なし道路延長率、歩道なし道路延長率、自転車歩行車道延長率は有意でなかった。このことから、自転車事故は暴露量の影響を最も受けやすく、自転車通行施設では事故リスクはほとんど変化しないことが明らかになった。すなわち、歩道があっても、必ずしも安全であるとはいえないと考えられる。

表 6 自転車通行施設の比較結果

	係数	t 値
定数	-9.339	-12.093 **
自転車交通量	0.431	11.882 **
車両交通量	0.806	10.032 **
区間延長	0.470	6.375 **
歩道なし延長率		
歩道あり延長率	0.144	0.702
自転車歩行車道延長率	0.250	1.134
log likelihood		-1170.8
AIC		3.846
過剰分散パラメータ ϕ		3.960

5.2 事故類型の比較

次に、事故類型別の事故リスクの比較を行う。自転車事故のリスク変化要因を把握するために、4.1 のパターン③に交差点条件、通行施設、走行条件を表す変数を追加し、分析を行った。なお、事故類型による比較をするために、「車両×車両」と「二輪×車両」の事故類型についても同様に分析を行った。その結果 (表 7) と考察を以下に示す。

t 値に注目すると、5.1 と同様に、いずれの事故類型に関しても、暴露量の有意性が高かった。

暴露量の次に高い有意性を示したものは交差点密度に関する変数だった。各事故類型の交差点に関する変数の t 値に注目すると、「自転車×車両」では信号なし交差点密度 ($5.5m \leq W$)、「車両×車両」では信号交差点密度 ($5.5m \leq W$)、「二輪×車両」では信号交差点密度 ($W < 5.5m$) が一番高く、交差点の中で一番大きな影響を与えている。よって、「自転車×車両」事故は、他の事故類型と比較して、信号なし交差点での事故が多いと示唆できる。このことから、自転車事故は他の事故類型と比較して、信号なし交差点での出会い頭の事故が多いと考えられる。

また、これらの結果は、3 における基礎統計から得られた結果と比較すると、同様の結果であった。

表 7 事故類型別リスク分析結果

	自転車×車両		車両×車両		二輪×車両		
	係数	t 値	係数	t 値	係数	t 値	
定数	-10.7	-7.99 **	-8.67	-9.06 **	-10.2	-2.59 **	
区間延長	0.90	8.01 **	0.73	8.67 **	1.04	3.38 **	
自転車交通量	0.61	10.33 **					
二輪車交通量					0.89	4.44 **	
車両交通量	0.70	5.49 **	1.08	11.97 **	0.53	1.24	
信号	W \geq 5.5m	0.11	1.69	0.13	2.43 *	0.34	1.50
	交差点密度 W $<$ 5.5m	0.14	1.41	0.15	1.77	0.40	1.68
信号なし	W \geq 5.5m	0.09	2.27 *	0.05	1.62	0.05	0.37
	交差点密度 W $<$ 5.5m	0.03	1.16	0.02	1.14	0.23	1.19
歩道なし道路延長率	-0.04	-0.13	-0.21	-0.95	0.08	0.48	
歩道あり道路延長率							
路肩幅員	W $<$ 1	-0.30	-1.11	-0.24	-1.06	-0.55	-0.73
	1 \leq W $<$ 2	-0.17	-0.65	-0.11	-0.49	-0.30	-0.46
	2 \leq W						
歩道幅員	W=0	0.30	0.46				
	0 $<$ W \leq 2	-0.30	-1.47				
	2 $<$ W \leq 3	0.12	0.65				
	3 $<$ W						
法定最高速度	v $<$ 40	0.55	1.56	0.32	1.19	0.01	0.02
	v=50	0.41	1.30	0.06	0.25	0.06	0.07
	v=60						
車道幅員	W \leq 7	0.23	0.80	-0.18	-0.84	0.10	0.17
	7 $<$ W \leq 8	-0.04	-0.15	-0.14	-0.72	0.13	0.24
	8 $<$ W \leq 11	0.32	1.36	0.16	0.84	0.15	0.27
	11 $<$ W						
log likelihood		-793.33		-1484.92		-222.26	
AIC		4.05		7.48		4.07	

5.3 事故リスクマップ作成

事故リスクの地域特徴を把握するために、5.2のモデル式を用いて、各道路リンクの事故リスクを算出し、大小5段階に分類し、マップ化を行った(図8)。

マップ図より、神戸、阪神地区といった交通量の多い地域の事故リスクも高いが、宍粟市、神河町といった交通量の少ない地域でも事故リスクが高い地域も見られた。これより、自転車利用の少ない路線であっても道路の潜在的危険は高く、事故対策の優先順位は低くない地域もあることが示された。

6. まとめと今後の課題

本研究では、一般化線形モデルの有用性と自転車事故に影響を及ぼす要因を明らかにすることを目的として、道路リンクを単位とした事故リスク分析を行った。成果として以下のことがあげられる。

4より、一般化線形モデルの有用性が確認された。ま

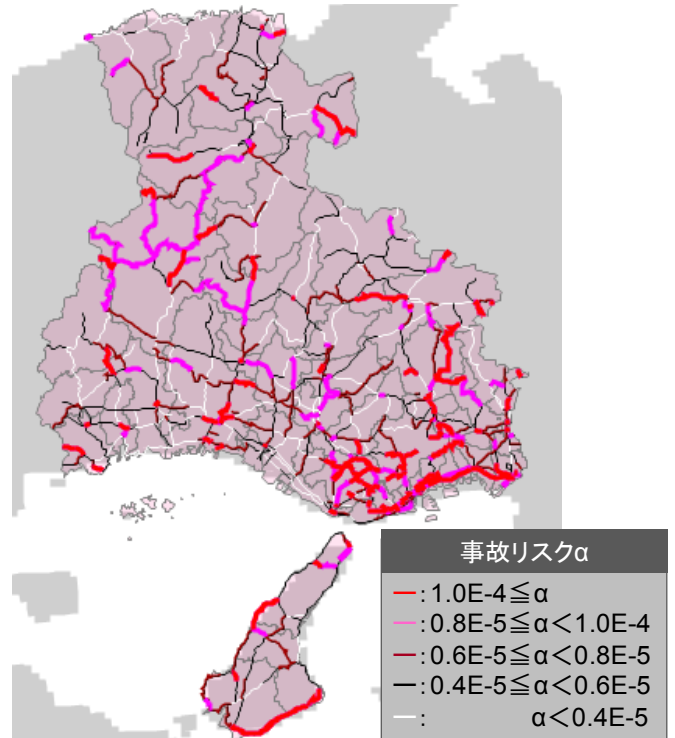


図8 兵庫県自転車事故リスクマップ

た、本研究のデータセットは常に過剰分散が発生していることが確認された。

また、5より、自転車事故は暴露量と交差点密度の影響を大きく受けることが統計的に示された。特に自転車×車両は信号なし交差点の影響を受けやすいことが統計的に示された。また、自転車通行施設の違いが事故リスクに与える影響は有意でないことから、歩道があっても必ずしも安全に繋がらないと考えられた。また分析結果から兵庫県下の事故リスクマップを作成し、地域特徴をつかんだ。

本研究の課題としては、本研究で用いたデータは道路交通センサスのリンク毎のデータであり、データに制限があることである。このデータが、区間延長が全て統一されたデータであったり、自転車通行施設単位のデータであったりすれば、過剰分散の影響をより小さくでき、より精度の増した分析が可能である。

参考文献

- 1) 村上功治：高知都市内交差点における事故リスク分析モデルの適用、第30回土木計画学講演集、2004
- 2) Ezra Hauer: Overdispersion in Modelling Accidents on Road Sections and in Empirical Bayesian Estimation, Accident Analysis and Prevention, Vol. 33, pp. 799-808, 2001.
- 3) 三浦利章、原田悦子：事故と安全の心理学、東京大学出版会、2008
- 4) 吉田伸一：自転車事故の現状と、自転車運転者の人的要因の分析、交通工学vol 40、2005

◆討議[東田淳 教授]

事故リスクは事故深刻度まで考慮するとあるが、どのように考慮したのか？

◆回答：分析において、被説明変数に事故件数を代入する際に、事故件数を「軽傷事故」と「死亡＋重傷事故」の2グループに分けて分析を行う。そして、その2つの分析結果を比較することにより、事故深刻度の違いを考慮した。

◆討議[東田淳 教授]

この研究は単年度だけの研究なのか？

◆回答：本研究はH17年の単年度を対象にした研究である。本研究で用いた道路交通センサデータは5年毎にしか入手できないので、単年度における分析となった。本研究の課題として、事故リスクの経年変化を確認することにより、分析結果の精度を向上させ、より詳細な考察が必要である。

◆討議[東田淳 教授]

本研究のアウトプットは何なのか？また、そこから導かれる自転車事故対策は何なのか？

◆回答：本研究のアウトプットの1つとしては事故リスクマップがあげられる。本研究で行った事故リスク分析結果の応用例として、事故リスクマップの作成を行った。また、本研究の目的は自転車事故の傾向を明らかにすることを主の目的としているため、自転車事故にどのような対策を施すかを提案することは研究の範囲外であり、行っていない。

◆討議[徳尾野徹 講師]

この分析は細街路にも適応できるのか？

◆回答：本研究で用いたH17道路交通センサでは、細街路のデータも存在するが、すべての道路区間において交通量が記載されているわけではなかった。本研究では、交通量と事故リスクを分けたモデル式により分析を行うことを前提としているので、本研究の対象外とした。もし、細街路の詳細な交通量データを入手することができれば、本研究の分析でも