

都道府県別の事故発生件数に関する統計的分析

2016SS021 金森敦司

指導教員：白石高章

1 はじめに

近年、運転マナーの悪さや交通事故に関するニュースが多く報道されている。私自身も自動車を運転している際に危険な場面に遭遇したことがあり、友人にも同じ体験をしている人がたくさんいた。このような報道や経験から交通事故に関心を持ち、インターネット等を利用して調べていると、交通事故に関するトラブルが多いとされる県が偏っているように感じた。そこで私は、全国の交通事故発生件数と道路交通量の関係を分析することによって、事故発生件数が多い県はどのような特徴があるか考察を行う。

2 データ

データは、平成 27 年度における各都道府県の交通事故発生件数 ([1]) と同年度の全国道路・街路交通情勢調査一般交通量調査集計整理表 ([2]) から引用した。各変数は、 x_1 (交通事故発生件数), x_2 (高速道路交通量), x_3 (一般国道 (DIO、商業地域) の交通量), x_4 (一般国道 (DIO、商業地域以外) の交通量), x_5 (一般国道 (平地部) の交通量), x_6 (一般国道 (山地部) の交通量), x_7 (主要地方道 (DIO、商業地域) の交通量), x_8 (主要地方道 (DIO、商業地域以外) の交通量), x_9 (主要地方道 (平地部) の交通量), x_{10} (主要地方道 (山地部) の交通量), x_{11} (一般都道府県道 (DIO、商業地域) の交通量), x_{12} (一般都道府県道 (DIO、商業地域以外) の交通量), x_{13} (一般都道府県道 (平地部) の交通量), x_{14} (一般都道府県道 (山地部) の交通量) としている。

分析を行うにあたって、「各都道府県の交通量」と「交通事故発生件数」を利用して分析をするはずだったが、都道府県ごとの道路の長さが異なったままでは適切な比較ができない。そこで、各都道府県の事故発生件数と道路交通量を道路延長距離で割ることにより、「1km 当たりの交通事故発生件数」と「1km 当たりの道路交通量」を用いて分析を行う。

3 分析方法

分析方法は、「因子分析」、「重回帰分析」、「重み付き重回帰分析」を用いた。分析は主に [3],[4],[5] を参考に行った。

4 因子分析

プロット図を分かりやすくするため各都道府県を北から南に $X1$ (北海道), $X2$ (青森県), $X3$ (岩手県) … $X47$ (沖縄県) とした。分析結果や P 値 から因子数を 4 と設定し、因子軸の回転方法はプロマックス回転で行った。バリマックス回転も行ったが、プロマックス回転と比較して解釈しにくいいためプロマックス回転で分析を行った。変数と因子の関係を把握し、因子得点を求めた。因子得点を図示したもの

が図 1 と図 2 である。

4.1 分析結果

第 1 因子は $X3$, $X11$, $X12$ の相関が高く、DIO や商業地域の道路を示している。関係が深い都道府県は 47 (沖縄県), 36 (徳島県), 31 (兵庫県) であることがわかった。47 (沖縄県) と 31 (兵庫県) は観光客数が上位であることがわかる。さらに、47 (沖縄県) では外国人観光客が、36 (徳島県) では日本人観光客が直近の何年かでかなり増えていることがわかった。このことが原因で人口集中地区の道路での事故が発生していると考察できる。

第 2 因子は $X1$, $X6$, $X2$, $X5$ の相関が高く、高速道路交通量が増えると事故発生件数が増えることを示している。関係が深い都道府県は 14 (神奈川県), 8 (山形県), 12 (栃木県) であることがわかった。3 県とも高速道路における事故が多かった。12 (栃木県), 8 (山形県) の因子得点が高いことは予想外だったが、高速道路で発生する事故は約 7 割が東北自動車道で発生しており、2 県とも含まれている。14 (神奈川県) は高速道路の交通量が非常に多いために因子得点が高いと考えられる。

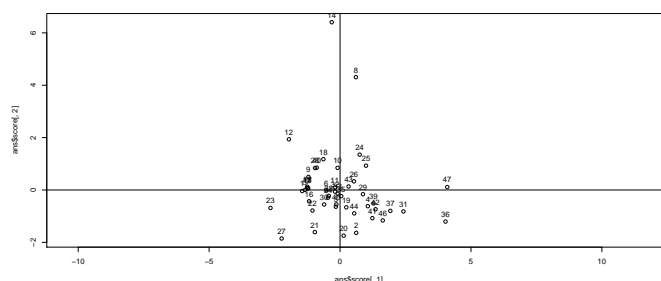


図 1 因子得点 1,2 因子

第 3 因子は $X14$, $X10$ の相関が高く、山地部の交通量を示している。関係が深い都道府県は 27 (大阪府), 23 (愛知県), 20 (石川県) であることがわかった。3 県の中でも特に関係が深い都道府県は 27 (大阪府) であり、大阪府の街作りについて調べてみたところ、人口集中地区の設備管理に時間と費用をかけていることがわかった。23 (愛知県) でもこのような傾向がみられ、これが原因で山地部を中心とした道路設備が整っていないと考えられる。

第 4 因子は $X13$, $X9$ の相関が高く、平地部の交通量を示している。関係が深い都道府県は 36 (徳島県), 41 (佐賀県), 29 (奈良県) であることがわかった。3 県の共通点を調べて

みた結果、可住地面積が狭い県であることがわかった。したがって、他県と比べて県民が住む場所がある程度決まっており、比較的暮らしやすい平地部がその対象になるのではないかと考えられる。

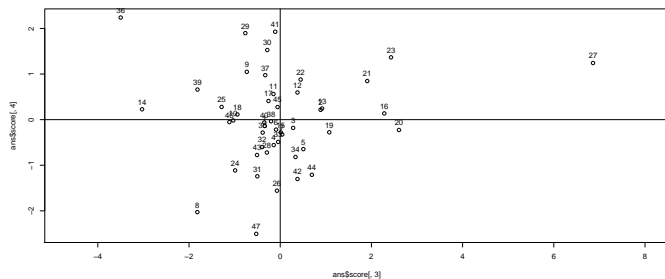


図2 因子得点 3,4 因子

結果となっている。重みを付ける前と比較すると、神奈川県が回帰診断図に含まれている。神奈川県は高速道路交通量が多く、やはり交通事故発生件数は高速道路交通量に深く関係していることがわかる。

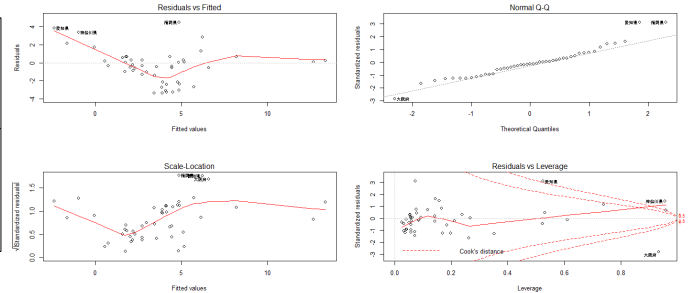


図3 回帰診断図

5 重み付き重回帰分析

交通事故発生件数を目的変数 y とし、各道路での 24 時間当たりの平均道路交通量をそれぞれ説明変数 $x_2 \sim x_{14}$ とし、重みとして新たに $weight$ をデータに追加して分析を行った。重みは各都道府県の人口を用いた。多重共線性が出ないようにステップワイズ法を利用して分析を行った。説明に不要な変数を除いた結果が表 1 である。

5.1 分析結果

表 1 重み付き重回帰分析の結果

変数	回帰係数	標準誤差	t 値	p 値
x_2	0.300	0.097	3.087	0.004
x_4	0.291	0.168	1.738	0.090
x_5	0.467	0.290	1.612	0.115
x_7	-0.477	0.208	-2.296	0.027
x_8	-0.845	0.298	-2.833	0.007
x_9	1.416	0.512	2.764	0.009
x_{10}	-1.790	1.381	-1.296	0.203
x_{13}	-1.588	0.575	-2.762	0.009
x_{14}	-1.784	1.090	-1.637	0.110

決定係数は 0.873, 自由度修正済み係数は 0.842 である。重みを付ける前と比べると値が高くなり、信頼性が増した。表 1 の結果から、有意性が高い変数は $x_2, x_7, x_8, x_9, x_{13}$ となった。その中でも回帰式に最も影響を与えている変数は x_2 である。また、回帰係数において正の影響を強く与えているのは x_9 , 負の影響を強く与えているのは x_{13} である。

5.2 回帰診断図の結果

図 3 の回帰診断図の結果から、愛知県、大阪府、福岡県、神奈川県の残差が大きく、予測より交通事故発生件数が多い

6 考察

本研究を行った結果、交通事故が起きる県は偏っているといえる。重み付き重回帰分析から神奈川県、愛知県、静岡県、大阪府、福岡県が他県よりも交通事故が多いことがわかった。交通事故発生件数の増加に強く影響を与える要因は x_2 (高速道路交通量), x_9 (主要地方道(平地部)の交通量) である。これらは重みを付ける前と同様であったが、決定係数が高くなり、神奈川県が新たに加わったことにより、 x_2 (高速道路交通量) がさらに交通事故発生件数の増加に強く影響を与える要因であることがわかる。 x_9 (主要地方道(平地部)の交通量) に関しては、人口集中地区でない平地部の道路は交通量が少なく、見通しが良いため速度超過を引き起こしやすいと考察できる。また、因子分析の結果から交通事故発生件数増加につながる要因として、各道路の交通量以外にも運転マナーの悪さや観光客の増加問題、道路設備による問題などが推測できた。

7 おわりに

本研究を通して事故の発生は交通量だけでなく、都道府県ごとに様々な要因があることがわかった。交通事故発生件数を減らしていくには各県に適した工夫が必要であるように思われる。

参考文献

- [1] 平成 27 年度における各都道府県の交通事故発生件数 <https://www.itarda.or.jp/materials/traffic/free>
- [2] 全国道路・街路交通情勢調査一般交通量調査集計整理表 <http://www.mlit.go.jp/road/census/h27/>
- [3] 金明哲:「R によるデータサイエンス データ解析の基礎から最新手法まで」。森北出版, 東京, 2007.
- [4] 中村永友:「R で学ぶデータサイエンス 2 多次元データ解析法」。共立出版, 東京, 2007.
- [5] 青木繁伸「R による統計的解析」, オーム社, 2009 年.