

携帯電話の位置情報を活用した 豪雪による市民生活への影響に関する研究

公共システム研究室 山口博哉

1. はじめに

近年、わが国では様々な地域で豪雪が見られ、豪雪の適切な対策が求められている。そのため、過去の豪雪においてどこでどれだけの支障が生じたのかを明らかにし、次の豪雪に備えて重点的な対応を要する箇所を明らかにすることが必要である。これまでは、インフラや生活サービスの供給機能については情報の把握がなされ、対策も多くの自治体で講じられていたが、市民の生活への影響を直接的かつ正確に把握することは困難であった。しかし近年では、携帯電話などの情報通信端末から利用者の位置情報データが記録されるようになり、これまでは断片的にしか把握することができなかつた人々の行動をある程度正確に把握することが可能となった。

そこで本研究では、携帯電話の位置情報により1時間ごとの人口分布を把握することができるNTTドコモが提供する「モバイル空間統計」を活用し、平成29年2月に鳥取県で発生した豪雪を対象として、通常時に比べ豪雪時に人々の行動がどこでどの程度異なっていたのかを定量的に明らかにする。具体的には、これらの時系列データから特徴的な部分を選び出す異常検知の手法をベースに、異常の程度や継続時間などのいくつかの側面から影響を評価するための指標を作成し、異常が生じる環境要因について統計的に明らかにする。その際、年代、居住地といった属性別での分析を行い、多角的に生活への影響を評価する。

2. 本研究の基本的な考え方

モバイル空間統計の時系列データから特徴的な部分を見出すため、異常部位検出を行なうとともに、その異常度に注目することで、人々への影響を定量的に評価する。その際、異常部位検出の手法の中で最も簡便な方法の一つである近傍法を用いる。ただし、この手法は変化の検知を目的としていることから、異常の程度や継続時間など、異常の様相を明らかにするものでは必ずしもない。そのため、これらの様相を定量化しうるよう指標を作成する。

その上で、全てのデータを用いた場合、居住地別のデータを用いた場合、年代別のデータを用いた場合の3パターンに分けて異常の評価を行う。居住地別による分析では、あるメッシュを対象に分析する際に、そのメッシュ内に居住している人とメッシュ外に居住している人を区分して分析

する。年代別による分析では20,30歳代、40,50歳代、60,70歳代の3つの年代に区分する。

3. モデルの構築

3.1 異常度の分析

異常部位を検出するにあたっては、部分時系列の集まりとして時系列データを表し、外れ値の検出問題として考える。各時刻の観測値をそのまま扱うのではなく、 m 個の隣接した観測値をまとめることにより、長さ τ の時系列データを m 次元ベクトルの集まりとしてまとめ直す。なお、この m の値は、算出した異常度が表す曲線が適度に滑らかになるように適当に決める必要がある。本研究では、用いたデータにおいて異常度が表す曲線が適度に滑らかになる $m=5$ を用いることとする。これにより、任意の地域 i における長さ τ の時系列データは、次式のように $N(=\tau-m+1)$ 個の m 次元ベクトル D に変換される。

$$D = \{x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(N)}\} \quad (1)$$

通常時の任意の日を $b(\in B)$ 、豪雪時の任意の日を $d(\in A)$ で表すと、豪雪時の t 時の行動と通常時の t 時の行動との距離は、次式で表される。ただし、データ x_i と距離 l_i の添え字 (d, t) は、 d 日の t 時が第一要素であるデータを表している。

$$l_i^{(d,t)} = \min_{b \in B} |x_i^{(d,t)} - x_i^{(b,t)}| \quad (2)$$

これにより、豪雪前の異常度 a_{i0} と豪雪後のそれ a_{i1} を次式により導出することができる。ただし、 A_0 は豪雪前の日、 A_1 は豪雪後の日の集合である。

$$a_{i0}(x^{(d,t)}) = m \ln l_i^{(d,t)} \quad (d \in A_0) \quad (3)$$

$$a_{i1}(x^{(d,t)}) = m \ln l_i^{(d,t)} \quad (d \in A_1) \quad (4)$$

3.2 異常の評価指標

豪雪前後の異常度の違いに着目し、異常の様相をいくつかの観点で数値化する。その一つ目として、豪雪前後における異常度の平均に着目し、一日当たりの影響を把握する。この視点に基づいた指標 μ_i は、次式で表される。ここに、 T は一日に

おける時間区分の数（一日あたりに 24 時間分のデータがあれば、 T は 24）である。

$$\mu_i = \sum_{d \in A_1} \sum_{t=1}^T \frac{a_{i1}(x_i^{(d,t)})}{|A_1|T} - \sum_{d \in A_0} \sum_{t=1}^T \frac{a_{i0}(x_i^{(d,t)})}{|A_0|T} \quad (5)$$

二つ目に、豪雪前後における異常度の最大値に着目し、最も異常が大きかった時点での影響を把握する。この視点に基づいた指標 λ_i は、次式で表される。

$$\lambda_i = \max_{d \in A_1, t} [a_{i1}(x_i^{(d,t)})] - \max_{d \in A_0, t} [a_{i0}(x_i^{(d,t)})] \quad (6)$$

なお、式(6)では瞬間的な異常の大きさを把握することができるが、その異常の継続性については不問であるのに対して、式(5)では継続した異常は一日あたりの大きさに平滑化して反映されているため、異常の大きさと継続性の双方を総合的に把握することができる。

異常の継続時間については、豪雪後のいつに通常時と同様の異常度に回復したのかで表すこととし、次式で求める。なお、 H^* はデータの収集期間である。

$$H_i = \frac{2\mu_i H^*}{\lambda_i} \quad (7)$$

3.3 地域の環境特性の影響に関する分析

式(5)～式(7)に示す異常の評価指標の大小は、各地域の環境特性に依存していると考えられる。そこで、これらの関係を回帰分析によって明らかにする。

4. 分析結果

異常と地域の環境特性に関する回帰分析の結果を表 1、表 2 に示す。なお、回帰分析の目的変数となる指標の値は、地域内の人口が豪雪時に増加しているものは正、減少しているものは負の値に変換する。これにより、回帰係数の正負は人口の増減にそれぞれ対応する。

40～50 歳代による分析の結果（表 1）に着目すると、全産業従業者数が負の値を示していることから、工場地帯や商業地域では、普段そこで働いている人が豪雪によって通勤できなくなり、その人口が減少することで異常が発生していると考えられる。人口総数が正の値を示していることから、人口総数が多くなる住宅街などでは普段外に出かけている人が豪雪によって外出を控えることにより、それらの地域で人口が増加することで異常が発生していると考えられる。また、混雑度大で係数が負の値であることから、交通量の

大きな幹線道路などで、豪雪の影響により車が一定時間の足止めを食らうことにより、通常時以下の交通量になったため異常が発生していると考えられる。60～70 歳代による分析の結果（表 2）に着目すると、全産業従業者数が負の値、人口総数が正の値を示していることは 40～50 歳代による分析と同様の理由と考えられる。スーパー総売り場面積、娯楽施設数が負の値を示していることから、高齢者が普段の外出先としている大きな商業施設や娯楽施設への外出を控え、近隣の小さなスーパーを利用、または、外出そのものを控えたため、その地域で人口が減少して異常が発生したと考えられる。

図 1 は、年代別の異常の評価指標の値が大きい地域をランキングし、そのランキングで上位 30 までに入った年代の割合を示したものである。この結果より、60、70 歳代の割合が高くなっており、異常が高齢者に発生しやすく、また、継続しやすいことが分かる。また、異常度の最大値に着目した評価指標 λ_i の結果では、20、30 歳代と 60、70 歳代は均等に上位に入っており、豪雪による異常が生じた時点では若者と高齢者で等しく異常が発生しやすいことが分かる。

表 1 40～50 歳代による分析結果（平均）

変数	偏回帰係数	P 値	トレランス	VIF
全産業従業者数	-0.0013	0.000**	0.89	1.12
人口総数	0.0009	0.000**	0.90	1.11
混雑度大	-0.2041	0.020*	0.94	1.07
定数項	0.0646	0.361		

修正 R2 : 0.39, * : P<0.05, ** : P<0.01

表 2 60～70 歳代による分析結果（平均）

変数	偏回帰係数	P 値	トレランス	VIF
全産業従業者数	-0.0018	0.00**	0.29	3.41
病院総面積	0.0001	0.00**	0.66	1.52
飲食店数	0.0756	0.03*	0.33	3.00
娯楽施設数	-0.6200	0.02*	0.68	1.47
定数項	1.0276	0.00**		

修正 R2 : 0.34, * : P<0.05, ** : P<0.01

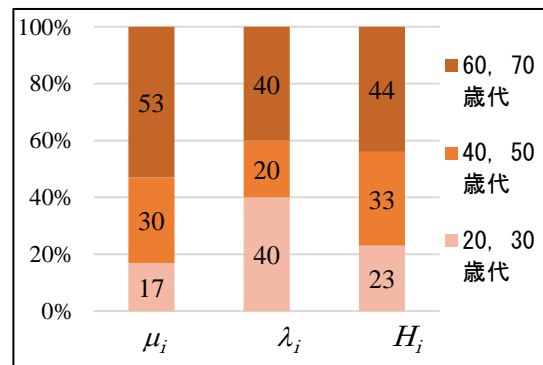


図 1 上位 30 までにおける各年代の割合