

# 空間統計モデルを用いた面補間法に基づく ODフローの空間詳細化

爲季 和樹<sup>1</sup>・堤 盛人<sup>2</sup>

<sup>1</sup>学生非会員 筑波大学 大学院システム情報工学研究科 (〒305-8573 茨城県つくば市天王台1-1-1)

E-mail: tamesue.kazuki@sk.tsukuba.ac.jp

<sup>2</sup>正会員 筑波大学教授 システム情報系社会工学域 (〒305-8573 茨城県つくば市天王台1-1-1)

E-mail: tsutsumi@sk.tsukuba.ac.jp

空間データの空間詳細化の方法論として面補間法が知られている。しかし面補間法が対象としているのは主に面データであり、ODフローを対象とした空間詳細化に関する研究は少ない。特にフローデータにおいては、四段階推定法の分布交通量の予測段階の様に、発生量と集中量が与えられた下で各ゾーンのフロー量を予測するケースが存在するなど、面データと異なりフローデータ特有の性質を考慮する必要がある。本研究ではフローデータの性質を考慮した上で、フローデータにおける空間詳細化のケースを二つに分類し、その両ケースにおける統計学的アプローチによる空間詳細化手法を重力モデルをベースに構築する。実証分析では我が国のODデータに適用し提案手法の精度の検証を行う。

**Key Words :** *origin-destination flow, spatial disaggregation, areal interpolation, spatial statistics, gravity model*

## 1. 序論

ある空間単位で集計されたデータを任意の空間単位での値に変換・補間する方法論として面補間法が知られている。面補間法における大前提として、補間前後の空間単位における変数の総和が一致するという体積保存則と呼ばれる性質を満たす必要がある点が挙げられる。例えば、都道府県単位で集計されたデータを、より空間的に詳細な市町村単位で予測した場合、その予測値の合計は元のデータの合計と一致していなければならない。面補間法の精緻化に関する研究は近年においても盛んであり、近年では空間的相関を考慮した統計モデルを援用し精緻化を行った研究もある (eg. 村上・堤, 2011)。

しかしながら、面補間法に関する研究は、主に面データを対象として発展してきており、フローデータを対象とした空間詳細化の研究はほぼ皆無である。また、フローデータはその性質上、点データや面データと比較して複雑であり、面データ等を対象として開発された今までの空間詳細化の方法論を単純にそのまま適用することは難しい。特に、面補間法における体積保存則の「体積」は、面データの場合は一意であり理解するのは容易だが、フローの場合は満たすべき体積が必ずしも一つのみではない (詳しくは第2章)。

面補間の概念をフローデータに適用した例としては、Boyle and Feng (2002) が挙げられるが、彼らの研究では大ゾーンでのデータを基に重力モデルのパラメータを推定し、得られたパラメータ推定値を用いて詳細なゾーンでのフロー量を外挿している。この方法では、体積保存則を満たすことは不可能であるため、得られた予測値の合計値が元のデータと整合するよう調整を行う、という極めてアドホックな方法をとっている。Boyle and Feng (2002) の様な外挿によって詳細な空間単位でのフロー量を予測する方法は、パラメトリックモデルにおいて仮定する確率分布の多くが有する再生性を無視しており統計学的に誤った方法である。以上のことから本研究では、面データとは異なるODフロー特有の性質を把握し、かつ統計学的理論に整合するフローデータの空間詳細化の方法論を提案することを目的とする。

本稿の流れは次の通りである。まず第2章で、本研究で対象とするフローの空間詳細化に関する概説を行い、面データと異なるフローデータ特有の性質について説明する。続いて第3章では、統計学的アプローチによりフローの空間詳細化を行うために本研究で構築する統計モデルとそのパラメータ推定法を説明する。さらに、第2章で触れたフロー特有の空間詳細化の性質を考慮した上で条件付き期待値を導出し、モデルを用いた予測方法を

提案する。そして第4章にて本研究の提案手法を我が国のODデータに適用して予測精度を検証し考察を行い、第5章ではまとめと今後の課題について述べる。

## 2. ODフローの空間詳細化の性質

### (1) 用いるデータ

空間的な二次元平面上の空間単位で集計された地域統計データにおいては、データの空間詳細化は面補間の研究でも良く知られており、例えば都道府県単位で得られた変数を1kmメッシュや市町村単位といったより細かい空間単位で補間・推計を行う。その際、補間前の空間単位（観測単位）における変数の総和と、補間後の空間単位（対象単位）における変数の総和が一致する体積保存則が満たされるようにする。一方でODフローの場合は、一つのフローに発地と着地という二つの空間単位が存在する。そのため上述のような空間詳細化において主な対象とされる面データと同様に二次元平面上においてその事象を把握・表現することは容易ではない。そこで以下では、OD表を用いることでデータの次元の複雑さを回避しつつ本研究で対象とするODフローの空間詳細化の概念について説明を行う。

ODフローにおける空間詳細化には、大きく分けて二つのケースが存在すると考えられる。一つは、ある空間単位間でのフローが観測された下で、より詳細な空間単位間のフローを補間するケースである（以下ケース1）。これは図1(a)の様に表すことができ、このとき体積保存則として対象となる体積はOD表における各マスで表されるフロー量である。したがってケース1は、OD表を二次元表面として考えその各マスを地域と見たてれば、OD表の面補間として捉えることも可能であり、一般的な面補間の方法論をフローの空間化詳細化に援用することは容易であると解釈できる。

もう一つのケースは、ある空間単位におけるフローの発生量と集中量が観測された下で、その空間単位間のフローを補間するケースである（以下ケース2）。これは図1(b)に示す通り、OD表の行和と列和が与えられた下で各マスのフロー量を補間するもので、例えば交通工学において有名な四段階推定法の分布交通量の予測段階や、人口移動における転入と転出人口データから地域間の人口移動を補間するケースが挙げられる。ケース1と同様に考えると、ケース2では、体積保存則で対象となる体積は発生量と集中量となる。つまり、保存則で満たすべき体積が二つ存在することになる。面データとは異なり体積が二変量となる特殊なケースとなるため、面データを対象としてきた面補間の方法論がそのまま適用することが困難であるのは明らかである。

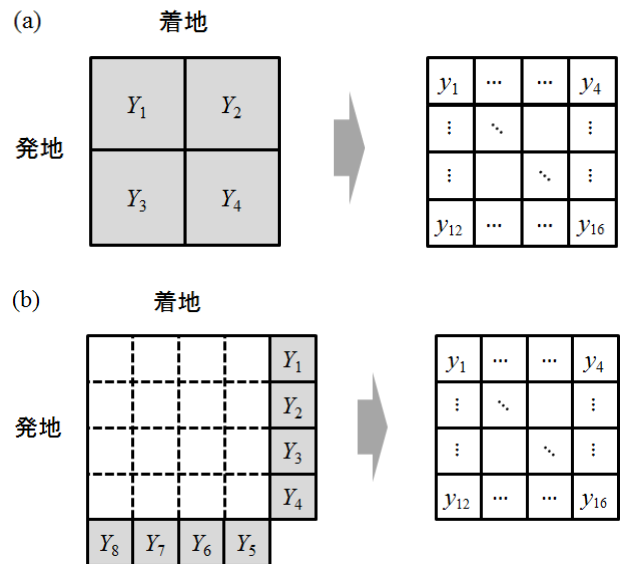


図1 ODフローの空間詳細化

(a) ケース1 (b) ケース2

## 3. 重力モデルを用いたフローの空間詳細化

### (1) モデル構築

本研究では、空間詳細化において対象とする空間単位間のODフローのデータ生成過程（DGP: Data Generating Process）としてフローがポアソン分布に従うと仮定し、ポアソン重力モデルを用いる。地域分析において古くから用いられている対数重力モデルでは、DGPとして対数正規分布を仮定しているが、交通トリップや人口移動のような観測値が台や人といった整数値しかとらないカウントデータに対して連続分布を仮定することは統計学的な誤りであることが古くから指摘されている（Flowerdew and Aitkin, 1982）。

今、観測されたフローを  $y = (y_1, \dots, y_n)$  としたとき、 $y$  がポアソン分布に従うと仮定すると、その確率密度関数は

$$\frac{\exp(-\lambda_i) \lambda_i^{y_i}}{y_i!} \quad (1)$$

と表すことができる。ここで  $\lambda_i$  はポアソン分布におけるパラメータであり、ポアソン重力モデルでは発地及び着地の規模を示す説明変数や発着地間の距離を表す説明変数を含む説明変数行列  $\mathbf{X}$  とその係数パラメータ  $\boldsymbol{\beta}$  を用いて

$$\lambda = \exp(\mathbf{X}\boldsymbol{\beta}) \quad (2)$$

としてモデルを構築する。ただし式(2)で表される様な古典的なポアソン回帰モデルでは、各フローが独立という強い仮定をおいている。ODフローも空間データのひとつであり、フローデータの空間的な自己相関を考慮

する必要性が近年認識されていることから (e.g., LeSage and Pace, 2008), 本研究では式 (3) の様に空間的なランダム効果を組み込んだ空間一般化線形混合モデル (SGLMM) を用いる:

$$\lambda_i = \exp(\mathbf{X}_i \boldsymbol{\beta} + S_i). \quad (3)$$

ここで  $\mathbf{S} = (S_1, \dots, S_n)$  は  $n \times 1$  のランダム効果ベクトルであり, 平均 0, 分散共分散行列  $\boldsymbol{\Sigma}$  の多変量正規分布に従う. 分散共分散行列  $\boldsymbol{\Sigma} = \sigma^2 \mathbf{R}(\rho)$  は地球統計学のクリギングにおける誤差項の分散共分散と同様に既知の相関関数 (共分散関数)  $\rho$  と  $\rho$  および  $\sigma^2$ , によって構造化される. ただし一般的なクリギングとは異なりナゲット効果は考慮しないものとする. また本研究では,  $\rho$  を所与のものとして固定することでモデルの簡素化を図った. 相関関数は, 観測データ間のユークリッド距離  $u$  に依存するとし, データ間の空間的な距離が近いほど相関性が高いことを明示的に考慮する. 関数の種類としては指数型やガウシアン型などが挙げられる (瀬谷・堤, 2014).

さて, 式 (3) に示されるように, 本研究では対象単位レベルでのフローのモデル化を行った. しかし, フローの空間詳細化にあたっては, 勿論対象単位でのフロー量を補間するのが目的であるため, 観測データ  $\mathbf{y}$  が存在しない. したがって, 式 (3) のままでは被説明変数を利用できないためモデルのパラメータを推定することは難しい. そこで, 第 2 章で述べたように, 対象単位でのデータのある程度集計した状態で観測されている, 観測単位でのデータ  $\mathbf{Y} = (Y_1, \dots, Y_N)$  を被説明変数として利用することを考える (ここで  $n > N$ ).

ケース 1 では, 観測単位のフロー  $\mathbf{Y}$  は対象単位のフロー  $\mathbf{y}$  を用いて式 (4) の様に表すことができる:

$$\mathbf{Y} = \mathbf{M}_1 \mathbf{y}. \quad (4)$$

ここで  $\mathbf{M}_1$  は  $N \times n$  の集計行列で, 要素  $(l, i)$  は  $y_i$  が  $Y_l$  に含まれていれば 1, それ以外では 0 の値をとる. つまり,  $y_i$  における発地  $j$  と着地  $k$  が双方とも  $Y_l$  における発地  $A$  及び着地  $B$  にそれぞれ含まれているときのみ要素  $(l, i)$  の値は 1 となる.

ポアソン分布は再生性を有するため,  $\lambda_i$  の平均を持つ互いに独立な複数のポアソン分布の和は,  $\lambda_i$  の総和を平均に持つポアソン分布に等しい. したがって, 式 (3) および式 (4) から, 観測単位間のフロー  $\mathbf{Y}$  は平均パラメータ  $\mathbf{A} = \mathbf{M}_1 \boldsymbol{\lambda}$  に従うポアソン分布となる.

$$\mathbf{A}_l = \mathbf{M}_{1l} \boldsymbol{\lambda} = \mathbf{M}_{1l} \exp(\mathbf{X} \boldsymbol{\beta} + \mathbf{S}) \quad (5)$$

ただし  $\mathbf{M}_{1l}$  は  $\mathbf{M}_1$  の  $l$  行である.

同様に, ケース 2 では, 対象単位でのフロー量  $\mathbf{y}$  と対象単位の発生量  $\mathbf{Y}_o$  及び集中量  $\mathbf{Y}_d$  の関係を, ケース 1 と同じ要領で集計行列  $\mathbf{M}_o$  と  $\mathbf{M}_d$  を定義すれば

$$\mathbf{Y}_o = \mathbf{M}_o \mathbf{y}, \mathbf{Y}_d = \mathbf{M}_d \mathbf{y} \quad (6)$$

と表され,  $\mathbf{Y}_o$  と  $\mathbf{Y}_d$  はそれぞれ平均パラメータ  $\mathbf{A}_o = \mathbf{M}_o \boldsymbol{\lambda}$  と  $\mathbf{A}_d = \mathbf{M}_d \boldsymbol{\lambda}$  に従うポアソン分布となる.

## (2) パラメータ推定法

式 (3) や式 (5) の様にランダム効果項を含む SGLMM のパラメータ推定法として MCMC を用いたベイズ推定が広く使われており, 例えば Christensen et al. (2002) や Paciorek (2007) などが詳しい.

パラメータ  $\boldsymbol{\beta}$ ,  $\mathbf{S}$ ,  $\sigma^2$  の事前分布は独立と仮定すれば, パラメータの事後確率密度関数は, 観測データ  $\mathbf{Y}$  の同時確率密度関数 (尤度関数)  $\pi(\mathbf{Y} | \boldsymbol{\beta}, \mathbf{S}, \sigma^2)$  とパラメータの事前確率密度関数の積に比例する. ただし  $\mathbf{S}$  は前節より  $\mathbf{S} \sim N(0, \boldsymbol{\Sigma})$  であるから,  $\mathbf{S}$  の事前確率密度関数は,  $\pi(\mathbf{S} | \sigma^2)$  と表すことができる.

$$\pi(\boldsymbol{\beta}, \mathbf{S}, \sigma^2 | \mathbf{Y}) \propto \pi(\mathbf{Y} | \boldsymbol{\beta}, \mathbf{S}) \pi(\mathbf{S} | \sigma^2) \pi(\boldsymbol{\beta}) \pi(\sigma^2) \quad (7)$$

よって各パラメータの条件付き事後分布は

$$\pi(\mathbf{S} | \dots) \propto \pi(\mathbf{Y} | \boldsymbol{\beta}, \mathbf{S}) \times \pi(\mathbf{S} | \sigma^2), \quad (8)$$

$$\pi(\boldsymbol{\beta} | \dots) \propto \pi(\mathbf{Y} | \boldsymbol{\beta}, \mathbf{S}) \times \pi(\boldsymbol{\beta}), \quad (9)$$

$$\pi(\sigma^2 | \dots) \propto \pi(\mathbf{S} | \sigma^2) \times \pi(\sigma^2), \quad (10)$$

となる. ここで,  $\boldsymbol{\beta}$  と  $\sigma^2$  の事前分布をそれぞれ  $\boldsymbol{\beta} \sim N(0, \boldsymbol{\Sigma}_\rho)$ ,  $\sigma^2 \sim LN(0, a)$  と設定すると, 式 (8) と式 (9) はポアソン分布と正規分布の合成となり,  $\mathbf{S}$  と  $\boldsymbol{\beta}$  の事後分布は標準的な分布にはならない. 事後分布からの直接の乱数生成を行うギブスサンプラーを適用することが困難であるため, MCMC におけるサンプリングでは代替法としてメトロポリス・ヘイスティングス (MH) アルゴリズムが必要となる. 本研究では MH アルゴリズムの中で代表的な酔歩サンプリングを用いることとする.

## (3) 条件付き期待値に基づく対象単位フロー量の補間

前節までで説明したモデルと推定法を用いて, パラメータの推定値と観測データが与えられた下での対象単位間のフロー量の条件付き期待値の導出を行う.

まずケース 1 に関しては, 第 2 章でも述べた様に体積保存則において満たすべき体積が単変量であることから, 空間詳細化の本質は面データのそれと変わらない. 今, 対象単位でのフロー量  $y_i \sim \text{Po}(\lambda_i)$  と,  $y_i$  を含む観測単位でのフロー量  $Y_l \sim \text{Po}(A_l)$  を考えたとき,  $Y_l$  が与えられた下での  $y_i$  の条件付き確率  $P(y_i | Y_l)$  は, 付録 (1) より二項分布

$$\text{bin}\left(Y_i, \frac{\lambda_i}{A_i}\right) \quad (11)$$

に従うので、条件付き期待値は

$$E\left[\text{bin}\left(Y_i, \frac{\lambda_i}{A_i}\right)\right] = Y_i \times \frac{\lambda_i}{A_i} \quad (12)$$

で求めることができる。Mugglin et al. (1999) ではこの条件付き期待値を用いて、ベイジアン面補間法を構築し面データでの空間詳細化を行っている。

ケース1と同様にケース2における条件付き期待値を導出するが、付録(2)の通りこれはケース1に比べて複雑である。対象単位の発生量  $Y_o \sim \text{Po}(A_o)$  および集中量  $Y_d \sim \text{Po}(A_d)$  が与えられた下での、対象単位でのフロー量  $y_i \sim \text{Po}(\lambda_i)$  の条件付き確率  $P(y_i | Y_o, Y_d)$  は、付録(2)より二項分布とポアソン分布の合成

$$\text{bin}\left(Y_o, \frac{\lambda_i}{A_o}\right) \text{bin}\left(Y_d, \frac{\lambda_i}{A_d}\right) \text{Po}(\lambda_i)^{-1} \quad (13)$$

によって表される。したがって条件付き確率は

$$E[y_i | Y_o, Y_d] = Y_o Y_d \frac{\lambda_i}{A_o A_d} \quad (14)$$

で求めることができる。

## 4. 実証分析

### (1) 使用するデータ

本章では、ケース2を対象として本研究の方法論を我が国のODデータに適用しその推定精度を検証する。本研究で使用するODデータは、平成22年国勢調査の移動人口集計における茨城県内の44市町村間人口移動である。このデータは、国勢調査で得られた全ての調査票を用いて平成17年と平成22年の常住地を比較することで、人口移動の転出入状況を市町村単位で集計したものである(総務省統計局, 2012)。転出入届に基づく業務統計である住民基本台帳人口移動報告のデータと異なり、本データは同一市町村内での移動も把握されているという特徴がある。このOD表の列和と行和をそれぞれ算出し観測データとみなし、第3章におけるモデル構築とパラメータ推定を行い、得られたパラメータ推定値を用いて式(14)の条件付き期待値を求める。

対象単位間フローの説明変数には、発地と着地それぞれにおける平成17年国勢調査の市町村別人口、発着地間の距離 ( $km^2$ )、そして地域内移動ダミーと、定数項を含めた5つを用いる。なお、発着地間の距離は各市町村の役所間のユークリッド距離として定義した。地域内距離には、Head and Mayer (2000), Redding and Venables (2004), Behrens et al. (2012) 等が用いられている近似距離指標

表1 パラメータ推定結果

	係数(平均値)	95%信頼区間	Geweke
定数項	-11.758	[-11.794, -11.721]	0.572
人口(発地)	0.904	[0.902, 0.905]	-2.528
人口(着地)	0.962	[0.961, 0.963]	-1.229
距離	-2.217	[-2.232, -2.202]	1.881
内々ダミー	1.580	[1.546, 1.614]	-0.758
$\sigma^2$	10.581	[9.912, 11.250]	0.310

$$d_{ii} = \frac{1}{3} \sqrt{\frac{\text{Area}_i}{\pi}} \quad (15)$$

で与えた。

### (2) 推定結果

モデルの推定にあたって、 $\beta$ の事前分布  $\beta \sim N(0, 10^6 \mathbf{I})$ 、 $\sigma^2$ の事前分布  $\sigma^2 \sim LN(0, 10^6)$  と設定し、 $\varphi = 20 km$  と固定した上でサンプリングの繰り返し回数を10,000回行い、そこからバーンイン1,000回を取り除いた9,000個のサンプルを推論として使用した。表1はそのパラメータ推定結果である。GewekeのZ値によるMCMCの収束診断では、人口(発地)の係数以外は5%検定において収束していると判断された。また、全てのパラメータの95%信頼区間において0が含まれないことから、各パラメータは有意であると判断できる。

対象単位でのフロー量の予測は、以上のサンプリングによって得られた各パラメータの平均値と標準偏差を用いて10,000回の乱数を発生させ、式(3)と式(6)に代入し式(14)の条件付き期待値を求めて10,000個のサンプルをとるシミュレーション方式により行った。

図2は横軸に市町村間人口移動の実測値、縦軸に予測値の10,000個のサンプルの平均値をとってプロットしたものである。図から明らかなように、大きな実測値で高い値のフローが予測ではほぼ0に近い値をとっており、予測精度が低いことが伺える。実測値として最も高い値は水戸市内の人口移動、次に高い値は日立市内の人口移

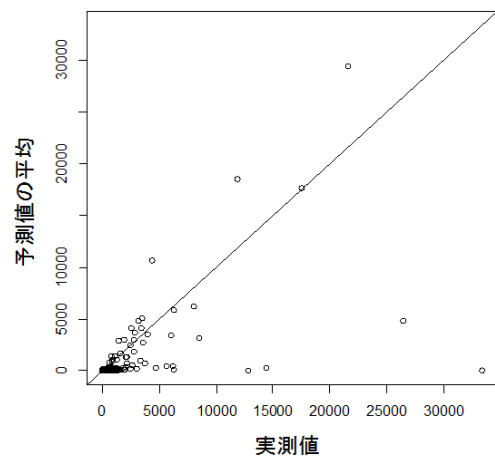


図2 フロー予測値の散布図

動, となっており, このことから地域内フローの予測精度が特に低いことが分かる. 予測精度の向上として内々フローを除外することも考えられるが, 空間詳細化の観点からは, そのような意図的なデータの除外は好ましくない. したがって, 同一モデル内にて内外フローの説明威力を保ちつつ, 同時に内々フローの異質性を考慮できるようなモデルの理論的開発が必要であると考え.

## 5. 結論

本研究は, ODフローにおける空間詳細化を, (i) ある空間単位間でのフローが観測された下で, より詳細な空間単位間のフローを補間するケース, (ii) ある空間単位におけるフローの発生量と集中度が観測された下で, その空間単位間のフローを補間するケース, の2ケースに大別し, それぞれのケースにおける補間手法をポアソン重力モデルを基に構築し, また条件付き期待値を導出した.

茨城県内市町村間人口移動データを用いた実証分析では全体の予測精度がい結果となったが, これはMCMCのアルゴリズムを安定化させるために, 使用した説明変数の数を少なくしたことでモデルの説明力低下を招いたことが原因と考えられる. 特に統計モデルを用いた面補間法では予測精度はモデル構築に大きく依存するため, 今後より一層のモデルの精緻化を行う必要がある. また, 特に内々フローの予測精度が悪かったことから, 内々フローの説明力が向上するようなモデルの開発を行うことが求められると考える.

その他の今後の課題としては, 推定法の精緻化が挙げられる. 本研究ではシンプルなMCMCアルゴリズムを採用したが, MHサンプリングにおける酔歩アルゴリズムのチューニングパラメータの調整や, Christensen and Waagepetersen (2002), Christensen et al. (2006), Chagneau et al. (2011) 等に見られるようにSGLMMのランダム効果項をメトロポリス・ランジュバン・アルゴリズムを用いて高速化するなど, アルゴリズムの改良の余地が十分残されている. また, パラメータ推定法として, マルコフ連鎖とEMアルゴリズムを組み合わせたMCEM (e.g., McCulloch, 1997) や, Gotway and Stroup (1997) の擬似尤度に基づく推定法などと比較・参考にして, パラメータ推定の安定化に努めたい.

謝辞: 本研究は JSPS 特別研究員奨励費 251786 の助成を受けたものである.

## 付録

### (1) ケース 1 における条件付き期待値の導出

確率変数  $X$  と  $Y$  が互いに独立で, それぞれポアソン分布  $Po(\lambda_1)$ ,  $Po(\lambda_2)$  に従うとき

条件付き確率  $P(X = x | X + Y = y)$  は

$$\begin{aligned} P(X = x | X + Y = y) &= P(X = x, X + Y = y) / P(X + Y = y) \\ &= P(X = x, Y = y - x) / P(X + Y = y) \\ &= P(X = x)P(Y = y - x) / P(X + Y = y) \\ &= \frac{e^{-\lambda_1} \lambda_1^x}{x!} \frac{e^{-\lambda_2} \lambda_2^{(y-x)}}{(y-x)!} \bigg/ \frac{e^{-(\lambda_1+\lambda_2)} (\lambda_1 + \lambda_2)^y}{y!} \\ &= \frac{y!}{x!(y-x)!} \frac{\lambda_1^x \lambda_2^{(y-x)}}{(\lambda_1 + \lambda_2)^y} \\ &= \binom{y}{x} \left( \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^x \left( \frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{(y-x)} \\ &= \text{bin} \left( y, \frac{\lambda_1}{\lambda_1 + \lambda_2} \right). \end{aligned}$$

したがってその条件付き期待値は

$$E \left[ \text{bin} \left( y, \frac{\lambda_1}{\lambda_1 + \lambda_2} \right) \right] = y \times \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

となる.

### (2) ケース 2 における条件付き期待値の導出

確率変数  $X$ ,  $Y$ ,  $Z$  が互いに独立で, それぞれポアソン分布  $Po(\lambda_1)$ ,  $Po(\lambda_2)$ ,  $Po(\lambda_3)$  に従うとき

条件付き確率  $P(X = x | X + Y = y, X + Z = z)$  は

$$\begin{aligned} P(X = x | X + Y = y, X + Z = z) &= P(X = x, X + Y = y, X + Z = z) / P(X + Y = y, X + Z = z) \\ &= P(X = x, Y = y - x, Z = z - x) / P(X + Y = y, X + Z = z) \\ &= P(X = x)P(Y = y - x)P(Z = z - x) / P(X + Y = y)P(X + Z = z) \\ &= \frac{e^{-\lambda_1} \lambda_1^x}{x!} \frac{e^{-\lambda_2} \lambda_2^{(y-x)}}{(y-x)!} \frac{e^{-\lambda_3} \lambda_3^{(z-x)}}{(z-x)!} \bigg/ \end{aligned}$$

$$\frac{e^{-(\lambda_1+\lambda_2)} (\lambda_1 + \lambda_2)^y}{y!} \frac{e^{-(\lambda_1+\lambda_3)} (\lambda_1 + \lambda_3)^z}{z!}$$

$$= \frac{1}{e^{-\lambda_1}} \frac{y!z!}{x!(y-x)!(z-x)!} \frac{\lambda_1^x \lambda_2^{(y-x)} \lambda_3^{(z-x)}}{(\lambda_1 + \lambda_2)^y (\lambda_1 + \lambda_3)^z}$$

第二項に  $x!/x!$  を掛け、第三項に  $\lambda_1^x/\lambda_1^x$  を掛けると

$$= \frac{1}{e^{-\lambda_1}} \frac{x!}{1} \left\{ \frac{y!}{x!(y-x)!} \frac{z!}{x!(z-x)!} \right\} \frac{\lambda_1^x \lambda_2^{(y-x)} \lambda_3^{(z-x)}}{(\lambda_1 + \lambda_2)^y (\lambda_1 + \lambda_3)^z}$$

$$= \frac{1}{e^{-\lambda_1}} \frac{x!}{1} \frac{1}{\lambda_1^x} \left\{ \frac{y!}{x!(y-x)!} \frac{z!}{x!(z-x)!} \right\} \times \left\{ \frac{\lambda_1^x \lambda_2^{(y-x)}}{(\lambda_1 + \lambda_2)^y} \frac{\lambda_1^x \lambda_3^{(z-x)}}{(\lambda_1 + \lambda_3)^z} \right\}$$

$$= \frac{x!}{e^{-\lambda_1} \lambda_1^x} \left\{ \binom{y}{x} \binom{z}{x} \right\} \left\{ \left( \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^x \left( \frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{(y-x)} \times \left( \frac{\lambda_1}{\lambda_1 + \lambda_3} \right)^x \left( \frac{\lambda_3}{\lambda_1 + \lambda_3} \right)^{(z-x)} \right\}$$

$$= \frac{x!}{e^{-\lambda_1} \lambda_1^x} \binom{y}{x} \left( \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^x \left( \frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{(y-x)} \times \binom{z}{x} \left( \frac{\lambda_1}{\lambda_1 + \lambda_3} \right)^x \left( \frac{\lambda_3}{\lambda_1 + \lambda_3} \right)^{(z-x)}$$

$$= \text{bin} \left( y, \frac{\lambda_1}{\lambda_1 + \lambda_2} \right) \text{bin} \left( z, \frac{\lambda_1}{\lambda_1 + \lambda_3} \right) \text{Po}(\lambda_1)^{-1}.$$

したがってその条件付き期待値は

$$E \left[ \text{bin} \left( y, \frac{\lambda_1}{\lambda_1 + \lambda_2} \right) \text{bin} \left( z, \frac{\lambda_1}{\lambda_1 + \lambda_3} \right) \text{Po}(\lambda_1)^{-1} \right]$$

$$= y \frac{\lambda_1}{\lambda_1 + \lambda_2} \times z \frac{\lambda_1}{\lambda_1 + \lambda_3} / \lambda_1$$

$$= yz \frac{\lambda_1}{(\lambda_1 + \lambda_2)(\lambda_1 + \lambda_3)}$$

となる。

## 参考文献

- 1) Behrens, K., Ertur, C., and Koch, W.: ‘Dual’ gravity using spatial econometrics to control for multilateral resistance, *Journal of Applied Econometrics*, Vol.27, pp.773–794, 2012.
- 2) Boyle, P. and Feng, Z.: A method for integrating the 1981 and 1991 British census interaction data, *Computers, Environment and Urban Systems*, Vol.26, pp.241–256, 2002.
- 3) Chagneau, P., Mortier, F., Picard, N., and Bacro, J.: A hierarchical Bayesian model for spatial prediction of multivariate non-gaussian random fields, *Biometrics*, Vol.67, pp.97–105, 2011.
- 4) Christensen, O.F. and Waagepetersen, R.: Bayesian prediction of spatial count data using generalized linear mixed models, *Biometrics*, Vol.58, pp.280–286, 2002.
- 5) Christensen, O.F., Roberts G.O., and Skold M.: Robust Markov Chain Monte Carlo methods for spatial generalized linear mixed models, *Journal of Computational and Graphical Statistics*, Vol.15, pp.1–17, 2006.
- 6) Flowerdew, R. and Aitkin, M.: A method of fitting the gravity model based on the Poisson distribution, *Journal of Regional Science*, Vol.22, pp.191–202, 1982.
- 7) Gotway, C.A. and Stroup, W.W.: A generalized linear model approach to spatial data analysis and prediction, *Journal of Agricultural, Biological, and Environmental Statistics*, Vol.2, pp.157–178, 1997.
- 8) Head, K. and Mayer, T.: Non-Europe: the magnitude and causes of market fragmentation in the EU, *Weltwirtschaftliches Archiv*, Vol.136, pp.284–314, 2000.
- 9) LeSage, J.P. and Pace, R.K.: Spatial econometric modeling of origin-destination flows, *Journal of Regional Science*, Vol.48, pp.941–967, 2008.
- 10) McCulloch, C.E.: Maximum likelihood algorithm for generalized linear mixed models, *Journal of the American Statistical Association*, Vol.92, pp.162–170, 1997.
- 11) Mugglin, A.S., Carlin, B.P., Zhu, L., Conlon, E.: Bayesian areal interpolation, estimation, and smoothing: an inferential approach for geographic information systems, *Environment and Planning A*, Vol.31, pp.1337–1352, 1999.
- 12) Paciorek, C.J.: Computational techniques for spatial logistic regression with large data sets, *Computational Statistics & Data Analysis*, Vol.51, pp.3631–3653, 2007.
- 13) Rosenthal, J.S.: Optimal proposal distributions and adaptive MCMC. In Brooks S., Gelman A., Jones G.L., Meng X. (eds.) *Handbook of Markov Chain Monte Carlo*. Chapman & Hall/CRC, Florida, pp.93–112, 2011.
- 14) 瀬谷創, 堤盛人: 空間統計学—自然科学から人文・社会科学まで一, 朝倉書店, 2014.
- 15) 総務省統計局: 平成 22 年国勢調査 移動人口の男女・年齢等集計結果 結果の概要, <http://www.stat.go.jp/data/kokusei/2010/idou1/pdf/gaiyou.pdf>, 2012.
- 16) 村上大輔, 堤盛人: Kriging を用いた実用的な面補間法, *GIS—理論と応用*, Vol.19, pp.59–69, 2011.

(2014. 4. 25 受付)