



感染症拡大リスクの地理的分布の評価に関する研究
—新型コロナウイルス感染症の感染データの利用—

メタデータ	言語: jpn 出版者: 宮崎大学工学部 公開日: 2021-11-02 キーワード (Ja): キーワード (En): 作成者: 古井田, 健太郎, 小酒井, 雄也, 高塚, 佳代子, 岡崎, 直宣, 山場, 久昭, 油田, 健太郎, Koida, Kentaro, Kozakai, Yuya メールアドレス: 所属:
URL	http://hdl.handle.net/10458/00010282

感染症拡大リスクの地理的分布の評価に関する研究 —新型コロナウイルス感染症の感染データの利用—

古井田 健太郎^{a)}・小酒井 雄也^{b)}・高塚 佳代子^{c)}・岡崎 直宣^{d)}・山場 久昭^{e)}・油田 健太郎^{f)}

Study on Evaluation of the Geographical Distribution of Infectious Disease Spreading Risk —Usage of Infection Data of Covid-19—

Kentaro KOIDA, Yuya KOZAKAI, Kayoko TAKATSUKA,
Naonobu OKAZAKI, Hisaaki YAMABA, Kentaro ABURADA

Abstract

As epidemic prevention management for new coronavirus infection (COVID-19) infection, it is necessary to indicate the area where the infection is currently susceptible. However, it is difficult to quantitatively show the geographical distribution of potential infection risk and infectious disease spread risk. There is a foot-and-mouth disease infection simulation model “Keeling model” as a method that enables macro risk analysis of the spatial transmission process. The Keeling model models the probability of foot-and-mouth disease infection on each farm during the period of infection, and its population parameter λ is highly versatile in quantifying the susceptibility of each farm based on the infection status of surrounding farms. It is an expression. Therefore, in this study, we devised a method to index the susceptibility to human infectious diseases for each unit region, referring to the concept of indexing by the population parameter λ of the Keeling model. He also applied the proposed method to COVID-19 infection data and demonstrated its validity. The method of calculating the potential infection risk index newly proposed in this indexing is a value calculated only from the location and human factors of the unit area without using any individual infection data. Therefore, similar utilization effects are expected in the future outbreak of new infectious diseases. In addition, methods based on infection data developed to improve the accuracy of the potential infection risk index can also be fully applied. There is expected.

Keywords: Keeling model, COVID-19, Potential infection risk index, SIR model

1. はじめに

新型コロナウイルス感染症 (COVID-19) の防疫管理では、1) 迅速な感染拡大抑止、2) 経済活動の担保、3) 意思決定プロセスの明確化が重要であり、これら相反する課題のバランスを考慮する必要がある。1) に関しては、SIR モデルとして知られている Kermack and McKendrick モデルに基づくモデルベースの分析から「接触機会の 8 割減」や経験則から「3 密回避」といった施策が早くから打ち出され、一定の効果が発揮された。3) の観点では、医療現場の逼迫度を表す各種指数 (以降「逼迫度」と総称する) に基づく意思決定が有用とわかった。一方、2) のために実施されている GOTO トラベルキャンペー

ンは、開始当初 1) とのバランスを考慮して「東京のみを除外する」との意思決定がなされた。しかし、その頃、東京と同程度の逼迫度を示す地域は他にも存在しこの判断は医療現場の逼迫度だけでは説明のつかないものだった。これは、東京に潜在的に備わる感染症流行が拡大するリスク (以降“潜在的感染リスク”と記す) が、他地域より大きいゆえの感染症拡大リスクが大きいことを想定しての判断と推察される。しかし、潜在的感染リスクや感染症拡大リスクの地理的な分布を定量的に示すことは困難であるため、「東京除外」の明確な理由を示すことはできなかった。以上より、潜在的感染リスクや感染期間中の感染症拡大リスクの地理的な分布を明らかにする手法の開発には、上記 3) の観点において、大きな意義があると考えられる。

ところで、一般に疫学のアプローチは、何の分布に着目するかにより大きく 3 つに分類される。具体的には、年齢・性別などの人の属性による分布、感染者数の時間的分布、及び感染者発生時の地理的分布に分類される¹⁾。特に感染者数の時間的分布を予測するための手法研究は盛んであり、中でも SIR モデルに基づく解析の手法の研究報告は枚挙にいとまがない²⁾。

^{a)}工学専攻機械・情報系コース大学院生

^{b)}情報システム工学科学部生

^{c)}教育研究支援技術センター技術専門職員

^{d)}情報システム工学科教授

^{e)}情報システム工学科助教

^{f)}情報システム工学科准教授

これに対し、地理的な分布に着目した空間伝染過程は解析的にモデル化することが難しいため、エージェントシミュレーションのための空間伝染モデルが近年提案されている³⁾。しかし「エージェント=個人」としてモデル化するため、地理的なリスク分布というマクロな視点でのリスク分析には適していない。一方、空間伝染過程のマクロなリスク分析が可能な手法として口蹄疫の感染シミュレーションモデル“Keelingモデル”⁴⁾がある。Keelingモデルは感染期間中の各農場の口蹄疫の感染確率をモデル化したものであるが、その母数 λ は各農場の感染し易さを周辺農場の感染状況に基づき定量化する汎用性の高い式となっている。そこで本研究の先行研究では、各農場の“感染し易さ”は λ で指標化可能と考え、これを“(ウイルスの)飛散指数”として指数化し、本指数ベースの防疫管理システムを開発した⁵⁾。また、母数 λ の考え方を応用し、各農場に元々備わっている感染し易さのリスクを“潜在的感染リスク”として指数化し、衛生管理システムの枠組みを考案した⁶⁾。なお、2010年の口蹄疫の実データを使用し、潜在的感染リスク指数の大きさと感染の有無との相関関係を感染源からの近さ別に調べたところ、本指数は感染時の感染し易さをそのまま表している可能性を示唆する結果が得られた。なお、Keelingモデルの λ の汎用性より、上記指数化の考え方は口蹄疫に留まらず人の感染症にも適用可能と考えられる。

2. 研究目的

そこで本研究では、Keelingモデルの母数 λ による指数化の考え方を参考にし、人の感染症の感染し易さ(以降“感染症拡大リスク”と同義とする)を単位領域ごとに指数化する手法を考案することを目的とする。

3. 提案方法

3.1 データ

国内のCOVID-19の感染者状況データ全般は各都道府県の公式ホームページで公表され、各感染者の感染判明日、年代、性別、居住地、管轄保健所等のデータが、国内初の感染データ(2020年1月16日)から現在まで全て閲覧可能である。本研究ではそれらデータを利用した。

3.2 Keelingモデルと先行研究

Keelingモデルは以下の通りである。

$$P_{inf,i}(t) = 1 - \exp(-\lambda_i(t)) \quad (1)$$

$$\lambda_i(t) = \sum_{j \in \text{infections}} C_{ij} N_i N_j h(r_{ij}) \quad (2)$$

$$h(r_{ij}) = h_0(1 + r/r_0)^\alpha \quad (3)$$

k1. $P_{inf,i}(t)$ は着目する感受性農場 i が感染初日を起点として t 日目に感染する確率を表す。

k2. $P_{inf,i}(t)$ の母数 λ を成す N_i は着目農場 i の保有動物数、 N_j は当該時点で感染動物が1頭以上存在する各農場 j の全動物数を表す。

k3. また、 C_{ij} は農場 i が農場 j から感染させられる確率(感染率)を表し、これは動物種の組み合わせで決まるパラメータである。

k4. h_{rij} は農場間での空間伝染の確率を決定する伝達関数の一つであり、関数の値は農場間距離と負の相関関係にある。

感染率 C_{ij} は、1匹の感染動物が平均何頭にうつすかを表す実効再生産数を牛に関して正規化したものであり、動物種の組み合わせによりその数値は大きく異なる。

先行研究では上記 λ_i を農場 i の感染し易さの指数と見なす提案をした。また、上記 k2 の λ_i を以下のように変更した λ'_i を農場 i の潜在的感染リスク指数と定義した。

k2'. λ'_i を成す N_i は着目農場 i の保有動物数、 N_j は着目エリア全体の各農場 j の全動物数を表す。

この潜在的感染リスクの指数化では、動物数のより多い農場がより近くに多く存在する農場ほどいざ感染が発生した際には感染症を拡大させる潜在的なリスクが多く備わっていることを一般化・数値化している。

なお、上記以外にも、農場への人や車の往来や小動物の侵入等、当該リスク指数に影響を及ぼす要素は多々あるが、いずれの事象も動物数や農場の密集度に比例して発生するというのが Keelingモデルでの考え方である。したがって、諸々の影響力も含めて指数 λ'_i で表されるという考えである。

3.3 感染症拡大リスクの指数化の考え方

感染人の感染症拡大リスクの指数化には先行研究の口蹄疫の感染し易さの指数 λ の考え方をそのまま適用できない。理由としてはまず、感染の有無の考え方が人の感染症と家畜感染症とは異なるためである。口蹄疫のような感染力の強い家畜伝染病の場合、1頭でも感染が確認された農場では全頭殺処分ということが法律で定められている。このため、家畜伝染病の感染の有無の判断は農場単位で行われ、1頭でも感染が確認された農場では全頭感染動物と見なされる。また、オリジナルの Keelingモデルには実効再生産数に基づき定義される感染率の時間による不確定変動を扱う仕組みがないが、人の感染症では感染率の時間変動を考慮する必要がある。特に未知の感染症を扱う場合、感染率の時間変動の想定が必須であることは、COVID-19の昨今の事例より明らかである。以上より、人の感染症の感染症拡大リスク指数は、新規に定義する必要がある。

一方、潜在的感染リスクの指数化では、Keelingモデルの“農場”を“単位領域”(都道府県、保健所管轄領域など)へ置き換えることにより、先行研究の λ' の考え方をそのまま適用できると考えられる。潜在的感染リスクの数値化においては、感染の有無に感知しないため、先に述べた感染の有無の定義付けの家畜と人との違いを考慮する必要がないためである。

ところで、先行研究において「潜在的感染リスク指数の大きさと感染動物累積総数との間には相関関係がある」という知見が得られている。この知見より、人の感染者累積総数と潜在的感染リスク指数との間にも相関関係のある可能性が考えられる(A)。また、この知見の中で着目している感染動物累積総数(即ち、

人の場合、感染者累積総数)の直近の数値こそが、各時点での感染しやすさをダイレクトに表す数値(B)と考えられる。そこで、本研究では、人の感染症の感染症拡大リスク指数は、潜在的感染リスク指数との相関があると仮定したうえで、これを直近の感染者累積総数に強くに相関するように定義することとした。

3.4 感染症拡大リスク指数の定式化

以上を踏まえ、人の感染症拡大リスク指数(以降これを“ λ ”と記す)の定式化のコンセプトを以下のように定義した。

$$\text{単位領域 } i, \text{ 期間 } t \text{ での感染症拡大リスク } \lambda_i(t) \quad (4)$$

$$= \text{単位領域 } i \text{ の潜在的感染リスク指数 } \lambda'_i \quad (5)$$

$$+ \text{時間 } t \text{ におけるリスク増分 } \Delta r_i(t) \quad (6)$$

$$+ \text{単位領域 } i \text{ に起因するリスク或はモデル化誤差 } \Delta e_i(t) \quad (7)$$

まず、本指数(上記(4))は、(5)、(6)、(7)の和として構成される。ここで、上記(5)の潜在的感染リスク指数の数値化の考え方は先に述べた通りである。即ち、3.2章の(2)式に関し、 k_2 を k_2' に置き換え、農場を単位領域に置き換えるという考え方で数値化できる。

次に(4)の感染症拡大リスク指数 $\lambda_i(t)$ の数値化に関しては、3.3章(B)で示した考え方(即ち、感染症拡大リスク指数は、直近の感染者累積総数に完全に相関するように定義するという考え方)に従うとし、このための一手法を考案した。即ち、各単位領域における $\lambda_i(t)$ の大小関係が直近の単位領域間での感染者累積総数の大小関係と完全に順序相関するように(C) $\lambda_i(t)$ の値を決定していくということである。具体的に、個々の $\lambda_i(t)$ を構成する $\lambda_i(t)'$ (5)は一意に決まる数値であるため、上記下線(C)で示す順序相関を実現させるうえで必要な最小の数値を、(6)と(7)の和($\Delta r_i(t) + \Delta e_i(t)$)として適宜付け加えるという方法を考案した。ところで、リスク増分の和 $\Delta r_i(t) + \Delta e_i(t)$ に関しては、もしこれを $\Delta r_i(t)$ と $\Delta e_i(t)$ に分割できるならば、上記(7)で定める意味より、 $\Delta e_i(t)$ は本来時間に依存しない定数(Δe_i)として扱われるべきものと考えられ、切り出された Δe_i は当該領域*i*に潜在的に備わるリスクの一部と見なすことが妥当となる。その場合、潜在的感染リスク指数は従来の λ'_i から $\lambda'_i + \Delta e_i$ へ修正されるのだが、その内訳は、 λ'_i はKeelingモデルの母数 λ_i に起因する潜在的感染リスクの数値であるのに対し、 Δe_i はKeelingモデルで説明されないその他全ての潜在的感染リスクを網羅する数値となる。

4. 検証実験

以上の感染症拡大リスク指数の定式化では、人の感染者累積総数と潜在的感染リスク指数との間に相関関係のある可能性を仮定していることは3.3章の下線(A)で述べた通りである。そこで、4.1章では、この仮定の正しさを検証した結果を示す。また、感染者数と潜在的感染リスク指数との間に相関関係についても併せて示す。更に、4.2章では、感染症拡大リスク指数の定式化の中で、従来の潜在的感染リスク指数 λ'_i を使用するよりも、新たに提案する潜在的感染リスク指数 $\lambda'_i +$

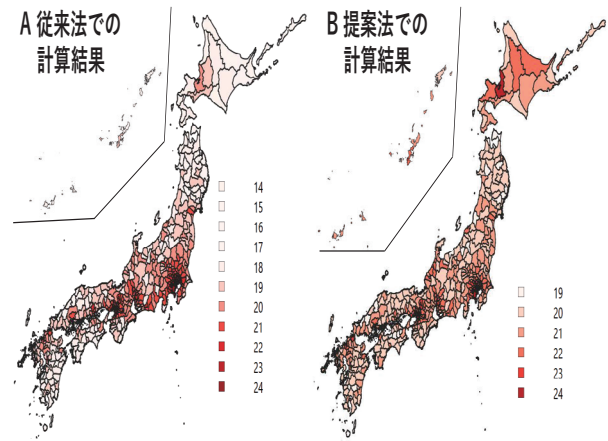


図 1. 潜在的感染リスク指数の地理的分布

Δe_i を使用する方が、感染症拡大リスク指数の精度が高くなることを実証的に示す。

4.1 従来法で算出される潜在的感染リスク指数の妥当性を検証するための実験(結果と考察)

国内のCOVID-19の2020年1月16日から12月31日までの感染者状況データ(3.1章)を利用し、感染者累積総数及び感染者数と潜在的感染リスク指数との相関関係を各々調べた。具体的な手順として、まず、指数化する上での単位領域を、今回は保健所管轄領域として設定した。そして、各単位領域*i*ごとの潜在的感染リスク指数 λ'_i を求めた。参考のために、計算結果の地理的分布を視覚化したものを図1-Aに示す。

更に、この指数の降順に単位領域名を列挙した。これを順序関係(1)とする。一方、感染発生時(2020/1/16)から各月末(3/31、4/30、…、12/31)までの感染者累積総数を単位領域ごとに求め、その降順に単位領域名を列挙し、これを順序関係(2)とする。そして、順序関係(1)と(2)の間の順序相関係数を、母集団の分布を仮定しない(つまり、データの順位しか分かっていない場合でも利用可能な)ノンパラメトリック検定であるスピアマン順序相関係数を算出する手法⁷⁾で求めた。相関係数の計算結果のグラフを図2-1に示す。また、潜在的感染リスク指数と各月の感染者数との相関係数も同様の方法で求めた。計算結果のグラフを図2-2に示す。グラフの横軸は時間、縦軸は相関係数を表している。グラフが示す相関係数は、いずれのグラフにおいても、概ね0.7より大きな値を示している。スピアマン順序相関係数 r の見方として、 $0.2 < r \leq 0.4$ は“低い相関あり”、 $0.4 < r \leq 0.7$ は“相関あり”、 $0.7 < r < 1$ は“高い相関あり”と評価されることから、本結果から次のことが言える。

- 先に仮定した通り、感染者累積総数と潜在的感染リスク指数との間には、概ね高い相関関係が示された。また、感染が進むにつれ徐々に相関が高まっていくことも確認された。
- また、感染者数と潜在的感染リスク指数の間にも、概ね高い相関関係が示された。このことから、感染症拡大リスクの地理的分布を明らかにするうえで、従来の潜在的感染リスク指数は利用可能と考えられる。

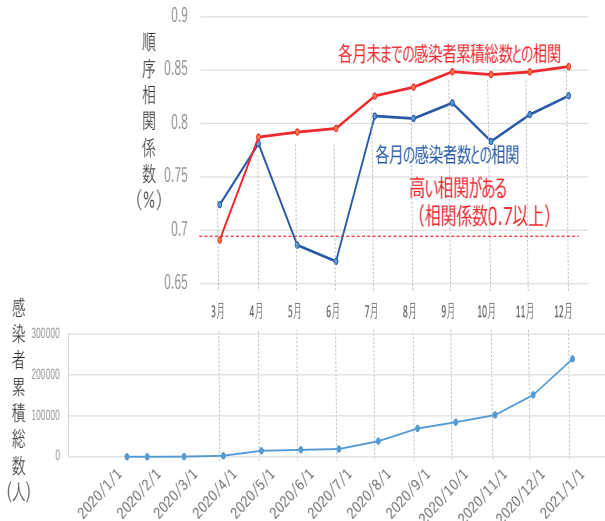


図 2. 従来法での潜在的感染リスク指数との相関関係

4.2 提案法で算出される潜在的感染リスク指数の有用性を検証するための実験（結果と考察）

新たに提案する潜在的感染リスク指数 $\lambda'_i + \Delta e_i$ と月々の感染者数と潜在的感染リスク指数との相関関係を調べた。ただし、今回は暫定的に、各单位領域 i の時間に依存しないリスク増分 Δe_i 、 $(\Delta r_i(t) + \Delta e_i(t))$ は、各月までの平均値として定めた。そこで、 M 月までの平均によるリスク増分 Δe_i を Δe_i^M と表記するようにする。参考のために、計算結果の地理的分布を視覚化したものを図 1-B に示す。結果のグラフを図 3 に示す。グラフの横軸と縦軸の意味は図 2 と同じである。各グラフの見方は、例えば薄いブルーのグラフは、4 月末時点で得られる潜在的感染リスク指数 $\lambda'_i + \Delta e_i^4$ と各月 (3 月～12 月) の感染者数との順序相関を表している。また、5 月末時点で得られる潜在的感染リスク指数 $\lambda'_i + \Delta e_i^5$ と各月の感染者数との順序相関を表し、以降同様に、潜在的感染リスク指数 $\Delta e_i^M (M=4, \dots, 12)$ と各月の感染者数との順序相関を表している。また、図 3 の結果の中で最も精度の良いのは潜在的感染リスク指数 $\lambda'_i + \Delta e_i^{12}$ を使用した場合 (濃いピンクのグラフ) であり、この結果と従来法の潜在的感染リスク指数 λ'_i を使用した結果との比較のグラフが図 4 である。図 3, 図 4 から次のことが言える。

- 図 3 のように、新たな手法で算出された潜在的感染リスク指数 $\lambda'_i + \Delta e_i^M (M=4, \dots, 12)$ と月々の感染者数との間には、概ね高い相関関係が示された。
- また、 $\Delta e_i^M (M=4, \dots, 12)$ の M の値が大きくなっていくにつれ、相関関係は高まっていくことが確認される。このことは、感染期間が延び感染データが蓄積されていくにつれ、感染データから算出される Δe_i^M の精度が良くなっていくことを意味する。つまり、潜在的感染リスク指数 $\lambda'_i + \Delta e_i^M$ の新たな計算方法及び今回の Δe_i^M の設定方法は妥当と言える。
- このことを顕著に表すのが図 4 である。濃いピンクのグラフは今回算出された潜在的感染リスク指数 $\lambda'_i + \Delta e_i^M (M=4, \dots, 12)$ のなかで最も精度の良い 12 月までの感染データを使用して得られた潜在的感染リスク指

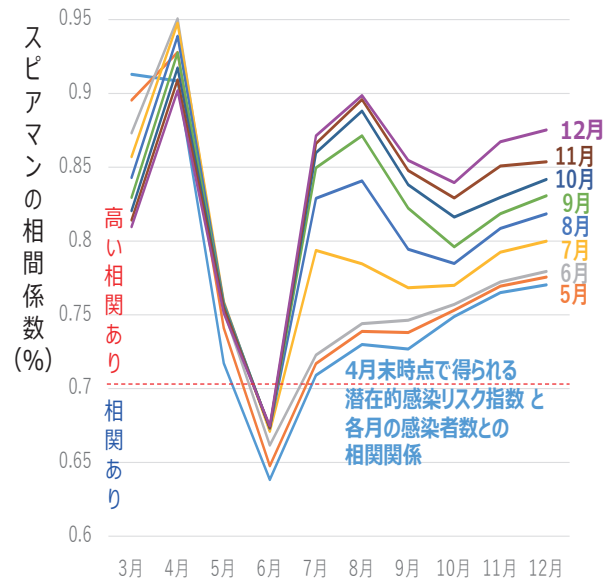


図 3. 提案法による潜在的感染リスク指数と月別感染者数との相関関係

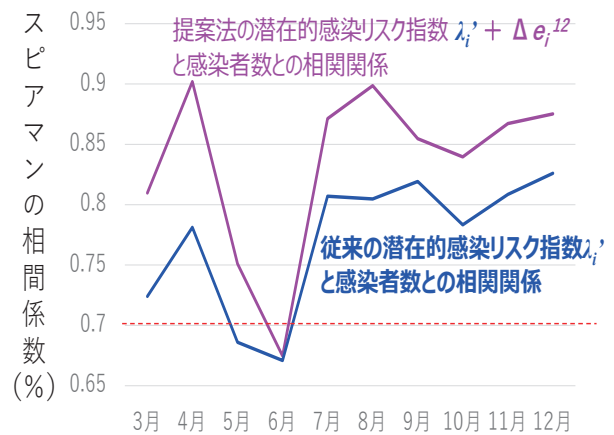


図 4. 潜在的感染リスク指数の従来法と提案法との精度の比較

数と感染者数との相関関係を表すグラフである。これに対し、ブルーのグラフは図 2 のブルーと同じであり、従来法で算出された潜在的感染リスク指数と感染者数との相関関係のグラフである。明らかに提案手法で得られた潜在的感染リスク手法との相関が高いことがわかる。

- ただし、5 月と 6 月の感染者数との相関は、従来法、提案法、双方の結果において顕著に低くなっている。このことは、2020 年 4 月 7 日より全国規模で相次いで出された緊急事態宣言による極端な防疫策の影響と考えられる。つまり、極端な防疫策の下では、潜在的な感染リスクの地理的分布のバランスが一時的に崩れる。5 月、6 月の相関関係の低下は、その結果の表れと考えられる。

5. まとめ

感染症拡大リスクを単位領域ごとに指数化する手法を考案した。また、提案手法を COVID-19 の感染データに適用し、その妥当性を示した。なお、今回の指数化で新たに提案した

潜在的感染リスク指数の算出方法だが、そのコアな部分は個別の感染データを一切使わず単位領域の場所的人的要因のみから算出される値であることから、将来の新たな感染症発生の際にも同様の利用効果が期待される。また、その潜在的感染リスク指数の精度を向上させるために開発した感染データに基づく手法も、十分応用できることが期待される。

参考文献

- 1) 蓑輪 眞澄: 流行調査, available from (http://idsc.nih.go.jp/training/9kanri/14_minowa.html), IDEC 感染症危機管理研修会資料, (2021/02 アクセス)
- 2) Sercan O. Arık, et al.; Interpretable Sequence Learning for COVID-19 Forecasting, NeurIPS, Submitted on 3 Aug 2020.
- 3) 石西、市川、他: エージェントベースシミュレーションによる高い致死性を持つ感染症来策におけるリスク分析手法の手尾杏, ISICIE, 27-7, 319/325, 2014.
- 4) Hayama, Y, et al.: Mathematical model of the 2010 FMD epidemic in Japan and evaluation of control measures, Preventive Veterinary Medicine, 112, 183/193, 2013.
- 5) Takatsuka, K, Sekiguchi S, et al.: “FMD-VS: A virtual sensor to index FMD virus scattering, PLOS ONE, <https://doi.org/10.1371/journal.pone.0237961> September 11, 2020.
- 6) 川添祐貴: 飼養衛生管理サポートシステムの開発, 宮崎大学工学部情報システム工学科卒業論文, 2019.
- 7) Spearman, C, et al.: The proof and measurement of association between two things, American Journal of Psychology, 72-101, 2307/1412159, 1904.