新しい小さいサンプルは古い大きいサンプルと同時に使うべきか：定数項の修正によるモデル更新の適用可能性

三古　展弘[1]

[1]正会員　神戸大学大学院准教授　経営学研究科（〒657-8501 神戸市灘区六甲台町2-1）
E-mail:sanko@kobe-u.ac.jp

定数項の修正は，サンプル数の多い古いデータで推定されたモデルの他のパラメータは移転可能という仮定のもと，最も移転性が低いと考えられる定数項（とスケールパラメータ）のみを新しい小さいサンプル数のデータで更新するものである．しかし，定数項を修正するよりも，新しい小さいサンプル数のデータのみを用いたほうが高い予測精度が得られるという報告もある．本研究では，古い時点と新しい時点のサンプル数がどの程度のときに，定数項の修正が新しい小さいサンプル数のデータのみを用いるよりも予測の改善に意義があるか検討する．分析の結果，いかなるサンプル数の組み合わせでも定数項の修正のほうが新しい小さいサンプル数のデータのみを用いたモデルよりも有意に良い予測をもたらすことはなかった．しかし，古い時点のサンプル数が大きく，新しい時点のサンプル数が小さいときには，統計的に有意ではないものの定数項の修正のほうが良い予測をもたらし，その予測値のばらつきが小さいことが明らかになった．

Key Words : *travel demand forecast, transferability, model updating, transfer scaling, data newness, number of observations, bootstrap*

1. はじめに

交通需要予測においては新しい時点のデータを用いることが重要である．新しい時点のデータを用いたモデルと古い時点のデータを用いたモデルを比較し，前者のほうがより良い予測をもたらすことが報告されている（Dissanayake et al., 2012[1]; Duffus et al., 1987[2]; Elmi et al., 1999[3]; Sanko, 2014[4]; Sanko et al., 2009[5]）．しかし，これらは十分なサンプル数のデータが得られている複数の時点でモデルを構築した場合に得られる知見である．ところが，実際には，古い時点のサンプル数は大きいが，新しい時点のサンプル数は小さい場合がある．

古い時点のサンプルは大きいが，新しい時点のサンプルは小さい場合は，次のようなときに発生すると考えられる．1つ目に，既に手元に多数のサンプルの古いデータと少数のサンプルの新しいデータがある場合である．2つ目に，手元にあるデータは古いがサンプル数は大きいことがある．ここでは，パーソントリップ調査のような大規模だが低頻度（三大都市圏では10年に1回）で実施される調査を想定している．このようなデータを用いて需要予測を行う場合，最新のデータでも10年以上前に収集されたものであることがある．このとき，既存の古いデータを使うか，新しいデータを独自に収集するかという判断が必要になる．しかし，大規模な調査を独自に行うことは困難であり，新たに行う調査から得られるサンプル数は小さいものにならざるを得ない．

サンプル数の大きい古いデータ，サンプル数の小さい新しいデータがあった場合，分析におけるデータの利用方法は次の3つが考えられる．

[1] サンプル数の大きい古いデータとサンプル数の小さい新しいデータを両方使用
[2] サンプル数の小さい新しいデータのみを使用
[3] サンプル数の大きい古いデータのみを使用

もし，1時点のデータしか用いないのであれば，[2]と[3]の比較が必要であり，既に筆者が報告している（三古[6]，Sanko[7]）．主な知見としては，新しい時点のサンプルが小さくても，古い時点の多数のサンプルのモデルよりも統計的に良い予測精度を持つ場合があることが確

認できた．しかし，これは言い方を変えると，新しい時点のサンプル数が小さければ，古い時点の多数のサンプルによるモデルよりも有意に良い予測をもたらさない場合があるということである．つまり，古い時点のデータを利用する意義がある可能性がある．

ここで，2つの時点のデータを用いた場合（上の[1]）を考える．これはモデルの更新として知られている．古い時点の多数のサンプルで構築したモデルを，新しい時点の少数のサンプルで更新する．つまり，古い時点のサンプルで構築したモデルを最小限の労力で収集した新しいデータによって少なくとも古い時点のデータのみを用いたモデルよりも改善しようという試みである．実際，モデルの更新は[3]よりも[1]のほうが良くなることを目的としているので，このことを実証した研究は多い．

本研究で対象とするのは上の[1]と[2]の比較である．モデルの更新は2つの時点のデータを両方用いているが，新しい時点の1つのデータを用いるだけでもモデルの更新よりもよい予測精度をもたらす場合もあるのではないかと考えられる．Badoe and Miller (1995)[8]とKarasmaa and Pursula (1997)[9]は新しい時点でのサンプルが少ない場合でも（例えば400～500サンプル），新しい時点に加えて古い時点のデータを利用することで予測精度の向上にはほとんど貢献しないと主張している．

今回対象とするモデルの更新法は定数項の修正である．これは，Atherton and Ben-Akiva[10]によって提案され，古い時点の多数のサンプルによって構築されたモデルの選択肢固有定数項（とスケールパラメータ）のみを新しい時点の少数のサンプルによって再推定し，それ以外の説明変数のパラメータは完全に移転可能とするものである．定数項の修正という方法は，非常に簡便にモデルを更新できる方法で理解もしやすいので，実務的にも利用価値が高いと考えられる．

本研究は実務的には次の問いに答えることができると考える．

1. 古い時点の多数のデータと新しい時点の少数のデータがあるとき，新しい時点のデータのみでモデルを構築するのが良いのか，古い時点のデータを用いたモデルを更新するのが良いのか．
2. 既存のデータがあるが，その後の交通状況の変化を反映するために新しいデータで分析したい．しかし，時間や費用などの制約から小規模な調査しか行えない．このとき，どのくらいのサンプルのときに新しい時点のデータのみでモデルを構築するのが良いのか，古い時点のデータを用いたモデルを更新するのが良いのか．

新しいデータを収集した場合，新しいデータのみを用いて良い予測ができるのであれば，それを古いデータと一緒に用いる必要はない．また，モデルの更新法を用いるのであれば，新しい時点のサンプルのみを使っても良い予測が行えるのであれば，必要以上にサンプルを抽出しているということになる．

2. 既存研究のレビュー

定数項の修正は，Atherton and Ben-Akiva[10]から引用すると，以下の動機によって提案された．

> ほとんどのモデルの特定化は，モデルによって明示的に説明されない要因を説明するために定数項を含んでいる．このような定数項の存在は，モデルが選択過程の総ての側面を捉えていないことを意味し，その捉えられていない要因は地域［時点］間で異なる可能性があるため，ある地域［時点］の定数項の推定値が別の地域［時点］で適当かは不明である．従って，時間，費用，所得，自動車利用可能性などの間の推定された関係が移転可能であるとする理論的根拠はあるが，定数項が移転可能であるという根拠はない．（［ ］内筆者．）

定数項の修正は，非常に簡便にモデルを更新できる方法で理解もしやすいので，実務的にも利用価値が高いと考えられる．ここで，これに関連した既存研究をいくつか紹介する．第1章で示した[1]と[3]を比較した既存研究は例えば，Sanko and Morikawa[11]に譲り，ここでは，[1]と[2]を比較した研究を紹介する．

Badoe and Miller[8]はカナダのトロントで1964年から得られた多数のサンプルと1986年から得られた少数のサンプルを用い，1964年のサンプル数は変化させずに1986年のサンプル数のみを変化させたときの1986年の予測精度を比較している．分析の対象はピーク時の出勤交通手段選択行動である．なお，この研究ではデータが2時点からしか得られていないため，2時点目の1986年のデータをモデル構築にも予測の検証にも用いているという問題がある．予測に用いるモデルは，定数項の修正を含む4種類のモデル更新手法を用いたモデルと1986年の少数のサンプルのみを利用したモデルである．その結果，1986年のサンプルが400から500程度あれば，1986年のデータのみを用いたモデルと，このデータに加えて1986年の多数のサンプルを用いたモデルの間に，予測精度にほとんど差がないという結論が得られている．

同様の分析をKarasmaa and Pursula[9]はフィンランドのヘルシンキの1981年と1988年のデータを用いて出勤時の交通手段・目的地選択モデルを対象に行い，類似の結論を得ている．

本研究の基本的な考え方は，2つの既存研究と類似しているが，以下のような工夫をし，より意味のある知見が得られるようにしている．

1. 既存研究では古い時点のサンプル数が固定されているため，古い時点のサンプル数が変わった場合については検討できない．本研究では古い時点と新しい時点の両方のサンプル数を変化させる．
2. 既存研究ではデータが2時点からしか得られていないため，2時点目のデータをモデルの構築にも予測の検証にも用いている．本研究では4時点のデータを使って3時点をモデルの構築に使用し，残った1時点をモデルの検証に使用する．
3. 既存研究では新しい時点での様々なサンプル数をランダムに抽出しているが，その抽出を1回しか行っておらず偶然性に左右される．本研究では，ブートストラップ法を用いて統計的に意味のある結論を導き出す．

3. データ

中京都市圏において1971年，1981年，1991年，2001年の4時点で得られた繰り返し断面データである，パーソントリップ(PT)調査データを用いる．モデルの構築には1971年，1981年，1991年のデータを用い，2001年のデータはモデルの予測精度の検証にのみ用いる．本研究で分析の対象とするのは鉄道，バス，自動車の3選択肢からの通勤交通手段選択行動である．データの詳細については三古[6]またはSanko[7]を参照されたい．なお，通勤の費用については通勤手当が支給されることが多いため考慮しない．

4. 方法論

多項ロジットモデルを例に説明するが，方法論は他のモデル構造の場合にも適用可能である．

(1) モデル

2つの時点として$t1$と$t2$を考え，[1] $t1$の時点で構築したモデルを$t2$の時点のデータを用いて更新した場合，[2] $t2$の時点のデータのみを用いてモデルを構築した場合を比較する．

ランダム効用理論に基づき，全効用を確定項と誤差項に分けて表現する．個人pの選択肢iに対する時点t（ここでは$t1$と$t2$を区別しないで定式化する）における効用関数の確定項V_{ip}^tを式(1)のように定式化する．

$$V_{ip}^t = \mu^t \left(\alpha_i^t + \sum_k \beta_{ik}^t x_{ikp}^t \right) \qquad (1)$$

ここに，μ^tは時点tのスケールパラメータ，α_i^tは時点tの選択肢iの選択肢固有定数項，x_{ikp}^tは時点tの個人pの選択肢iに対するk番目説明変数，β_{ik}^tはそれに対応するパラメータである．なお，スケールパラメータと変数のパラメータを識別することはできないので，スケールパラメータを1に固定する．

誤差項に独立で同一なばらつきを持つガンベル分布を仮定すると，時点tにおいて個人pが選択肢iを選択する確率P_{ip}^tは式(2)のロジット式で表現される．

$$P_{ip}^t = \frac{\exp\left(V_{ip}^t\right)}{\sum_j \exp\left(V_{jp}^t\right)} \qquad (2)$$

このとき，対数尤度関数は式(3)で表現され，これを最大化することによってパラメータを推定する．

$$L^t = \sum_p \sum_j y_{jp}^t \ln\left(P_{jp}^t\right) \qquad (3)$$

ここに，y_{jp}^tは時点tで個人pの選択結果が選択肢jであったとき1，そうではないとき0となるダミー変数．

[2]の$t2$の時点のデータのみを用いた場合は，上の議論で$t = t2$を代入する．

一方，[1]の$t1$の時点で構築したモデルを$t2$の時点のデータを用いて更新する場合，上の議論で$t = t1$を代入する．そして，$t2$のデータを用いて時点$t2$において個人pが選択肢iを選択する確率P_{ip}^{t2}は式(4)で表現される．

$$P_{ip}^{t2} = \frac{\exp\left(\mu^{t2} \left(\alpha_i^{t2} + \sum_k \left(\hat{\beta}_{ik}^{t1} x_{ikp}^{t2} \right) \right) \right)}{\sum_j \exp\left(\mu^{t2} \left(\alpha_j^{t2} + \sum_k \left(\hat{\beta}_{jk}^{t1} x_{jkp}^{t2} \right) \right) \right)} \qquad (4)$$

ここに，^は推定値を意味する．

ここで，μ^{t2}とα_i^{t2}のみを推定する．推定は式(3)で$t = t2$にすることで行うことができる．

モデルの予測精度は推定されたパラメータ$\boldsymbol{\beta}$，2001年の説明変数$\mathbf{x}$と選択結果$\mathbf{y}$を式(3)に代入することによって表現される．なお，スケールパラメータは[2]の場合には1，[1]の場合にはモデルの更新時に推定されたものを用いる．

(2) ブートストラップ[12]

まず，1971年，1981年，1991年のデータから通勤トリップをランダムに10000サンプルずつ抽出した．これは，各年における，そこからブートストラップを行うことになる，サンプル数の違いが結果に与える影響を避けるた

めである．また，10000サンプルとしたのは，ブートストラップにおける計算時間を節約するためである．また，予測対象年の2001年からもランダムに10000サンプルを抽出して検証に用いる．

ここで，5つの変数（y_1，y_2，n_1，n_2，b）を定義する．y_1とy_2 ($y_1 < y_2$)はデータ収集年，n_1とn_2 ($n_1 \geq n_2$)はそれぞれy_1，y_2において得られたデータのサンプル数，bはブートストラップの繰り返し回数である．ここで，上の2つの不等式は，古い時点y_1のサンプル数n_1で構築したモデルを新しい時点y_2の同じまたは小さいサンプル数n_2で更新することを意味している．

まず，それぞれのデータ年y (1971, 1981, 1991)においてサンプル数n (100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 10000の12通り)のデータを先に抽出した10000サンプルから200 (b = 1, 2, …, 200)回ランダムに復元抽出する．このとき，同一のyとbの組み合わせに対してnが小さいサンプルはnが大きいサンプルの一部分になるように抽出している．これによってy，nおよびbの組み合わせからなる3×12×200=7200通りのデータが生成された．

y_1，y_2，n_1，n_2について考えられる組み合わせは，yに関する3通りの組み合わせ((y_1, y_2) = (1971, 1981), (1971, 1991), and (1981, 1991))とnに関する78通りの組み合わせ(12×13/2)の3×78=234通り考えられる．この234通りのそれぞれについて200 (b = 1, 2, …, 200)回，[1] y_1時点のn_1サンプル（b回目抽出）によって推定された多項ロジットモデルをy_2時点のn_2サンプル（b回目抽出）によって更新したもの，[2] y_2時点のn_2サンプル（b回目抽出）によって推定した多項ロジットモデル，を計算した．[1]の計算をするのにy_1の時点のデータを用いて234×200=46800回，y_2の時点のデータを用いて更新するのに同じく46800回の合計93600回の推定を行う必要がある．[2]の計算をするのにはy_2の時点のデータを用いて46800回の推定を行う必要がある．つまり，延べ140400回の推定が必要になる．しかし，1971年のデータを用いてモデルを推定すればそれは1981年のデータで更新する場合にも1991年のデータで更新する場合にも使えるので，実際に作業をするときには140400回推定を行う必要はない．

予測精度は2001年のデータへの対数尤度で表現する．上の[1]と[2]の方法による2001年のデータへの対数尤度をそれぞれ$L1$ (y_1, y_2, n_1, n_2, b)と$L2$ (y_2, n_2, b)と表現する．

(3) 仮説検定

yとnに関して2つの場合，y_1とn_1（古い時点の多数のサンプル），y_2とn_2（新しい時点の少数のサンプル）を考える．ここで，$y_1 < y_2$かつ$n_1 \geq n_2$となるy_1，y_2，n_1，n_2について，以下の変数を定義する．

$$x_b = L2(y_2, n_2, b) - L1(y_1, y_2, n_1, n_2, b) \tag{5}$$

なお，ここで総てのb (= 1, 2, …, 200)に対して良好な推定結果が得られるわけではない．特にn_1やn_2が小さい場合においてこのことは問題となる．そこで，式(5)のx_bは，b回目のランダム抽出のときに$L1$と$L2$の両方が計算されたときにのみ，定義されることとする．

ここで，帰無仮説H_0と対立仮説H_1を下に示す．

$$H_0: x_b = 0$$
$$H_1: x_b \neq 0$$

ここで，次のzを定義する．

$$z = \frac{\overline{x_b}}{s(x_b)} \tag{6}$$

ここに，$\overline{x_b}$と$s(x_b)$はそれぞれ，x_bの平均と標準偏差．

ここで，x_bが標準正規分布に従っていると仮定すると，$z > 1.96$または$z < -1.96$で帰無仮説を棄却することができる．$z > 1.96$のときは新しい時点のサンプルのみを用いたほうが予測が有意に良いことを示し，$z < -1.96$のときは2つの時点のデータを用いる定数項の修正のほうが有意に良いことを示している．

5. 結果

(1) 推定結果

モデル化にあたって，男性ダミー（男性=1，女性=0），20歳以上ダミー（20歳以上=1，19歳以下=0），65歳以上ダミー（65歳以上=1，64歳以下=0），名古屋ダミー（名古屋市を出発地または到着地とする=1，そうではない=0）を定義した．モデルの変数の記述統計は三古[6]とSanko[7]に示されている．

各時点から抽出した10000サンプルのデータを用いてモデルを構築した結果を表―1に示す．1971年，1981年，1991年のデータを用いたモデル推定結果のほかに予測対象時点の2001年のデータを用いたモデルの推定結果も示す．モデルの解釈は三古[6]とSanko[7]を参照されたい．このとき，2001年のデータに適用して予測した予測精度も表―1に併記する（Log-likelihood on 2001 dataの行）．

また，表―1の1971年のモデルを1981年，1991年の10000サンプルのデータを用いて定数項とスケールパラメータを更新した場合，1981年のモデルを1991年の10000サンプルのデータを用いて更新した場合の，定数項とスケールパラメータの推定値のみを表―1に併記する．このとき，更新したモデルを2001年のデータに適用して予測した予測精度も表―1に併記する．

表─1　推定結果

Variables	1971 Est.	1971 *t*-stat.	1981 Est.	1981 *t*-stat.	1991 Est.	1991 *t*-stat.	2001[a] Est.	2001[a] *t*-stat.
Constant (B)	0.127	2.42	-0.392	-6.21	-0.638	-8.98	-1.03	-12.11
Constant (C)	-1.15	-9.84	-0.645	-4.65	0.301	1.96	0.560	2.23
Travel time [hr]	-0.606	-6.94	-1.81	-16.47	-1.59	-15.71	-2.60	-20.48
Male dummy (R)	0.577	8.59	0.787	8.70	0.812	7.53	0.511	3.89
Male dummy (C)	1.97	29.44	2.17	25.22	1.78	17.30	1.38	10.91
20 years old or older dummy (C)	0.900	8.28	0.764	5.78	0.776	5.18	0.511	2.06
65 years old or older dummy (B)	1.91	8.89	1.37	5.73	1.33	5.59	0.561	2.05
Nagoya dummy (C)	-1.12	-24.08	-1.77	-33.21	-2.18	-37.81	-2.21	-36.70
N (randomly drawn)		10000		10000		10000		10000
L(**β**)		-7776.86		-5985.02		-5300.58		-4716.28
L(**0**)		-8948.26		-8593.88		-8398.85		-8159.63
Adj rho-squared		0.130		0.303		0.368		0.421
Log-likelihood on 2001 data		-6521.95		-5225.15		-4801.79		Not applicable
Updated by 1981 data								
Constant (B)	-0.441	-11.91		–		–		–
Constant (C)	-0.986	-43.63		–		–		–
Scale parameter	1.27	44.49		–		–		–
N (randomly drawn)		10000		–		–		–
L(**β**)		-6138.88		–		–		–
Log-likelihood on 2001 data		-5672.04		–		–		–
Updated by 1991 data								
Constant (B)	-0.751	-16.18	-0.732	-13.26		–		–
Constant (C)	-0.667	-26.30	-0.271	-9.13		–		–
Scale parameter	1.21	42.41	1.02	45.29		–		–
N (randomly drawn)		10000		10000		–		–
L(**β**)		-5662.31		-5366.31		–		–
Log-likelihood on 2001 data		-5314.67		-4894.30		–		–

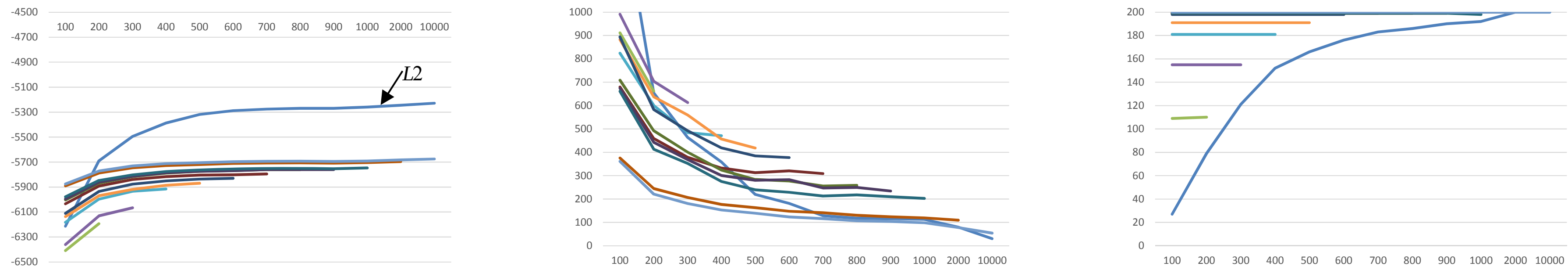
Note: (R), (B), and (C) notations refer to alternative-specific variables for rail, bus, and car, respectively. Variables without notations are generic.
[a] 2001 is the target year of forecast, and a model from 2001 is not required but is presented for a comparison purpose.

モデルの予測精度は，1971年のモデルよりも1971年のモデルを1981年のデータで更新したほうが良く，また1971年のモデルを1991年のデータで更新したほうがさらに良い．しかし，これは，1981年のデータや1991年のデータを単独で用いたときの予測精度には及ばない．つまり，新しい時点のデータのみを用いたほうが新しい時点のデータを古い時点のデータと一緒に用いるよりも高い予測精度があることが分かる．同様の議論が1981年のモデルを1991年のデータで更新した場合にも当てはまる．

(2)　予測精度の特徴

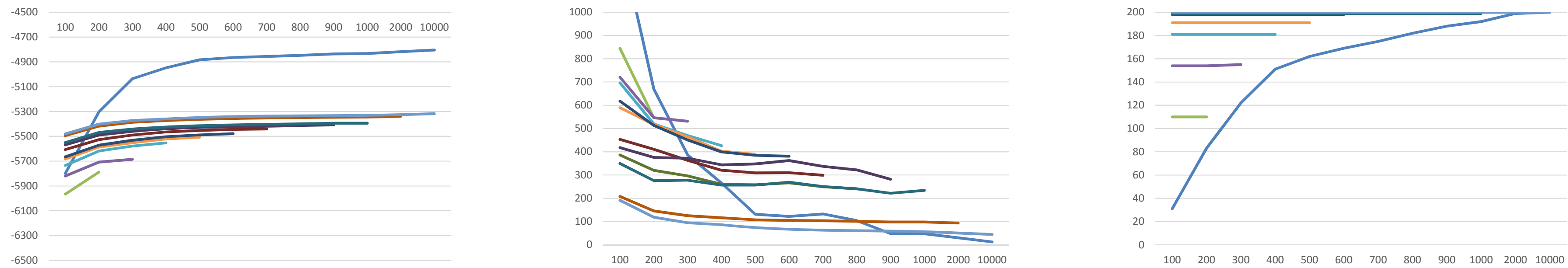
次にブートストラップによる結果を示す．まず，予測精度の平均値と標準偏差を考察する．先にも触れたが，n_1やn_2が小さい場合，モデルの推定に問題が発生することもある．ここでの平均値と標準偏差は$b = 1, 2, \ldots, 200$のうち，推定結果および予測結果が得られたもののみについて算出した．

$L1$ (y_1, y_2, n_1, n_2, b)と$L2$ (y_2, n_2, b)の平均値を図─1のパネル(a-1)～(a-3)に示す．この図の見方をy_1=1971，y_2=1981の場合のパネル(a-1)を例に説明する．横軸に新しい時点(y_2=1981)のサンプル数n_2をとって$L2$ (y_2=1981, n_2, b)の平均値を縦軸に描く．n_2は100から10000まで変化させているのでこの範囲で$L2$の平均値は得られている．今度は$L1$の結果についてn_1=100の場合について$L1$ (y_1=1971, y_2=1981, n_1=100, n_2, b)の平均値を縦軸に描く．ここで，$n_1 \geq n_2$なのでn_2=100の場合のみ描かれる．同様にn_1=200～10000の場合についても描く．これらの線については凡例がないので分かりづらいようにも思うが，線がn_2=100まで描かれているときはn_1=100のときの$L1$の平均値，線がn_2=200まで描かれているときはn_1=200のときの$L1$の平均値，というように考えると分かりやすい．なお，n_2=100～10000の範囲で描かれているのは$L2$ (y_2=1981, n_2, b)と$L1$ (y_1=1971, y_2=1981, n_1=10000, n_2, b)の2つである．どちらも青系統の色で描かれているが，濃いほうの図の中での触れ幅の大きいほうが$L2$ (y_2=1981, n_2, b)である．



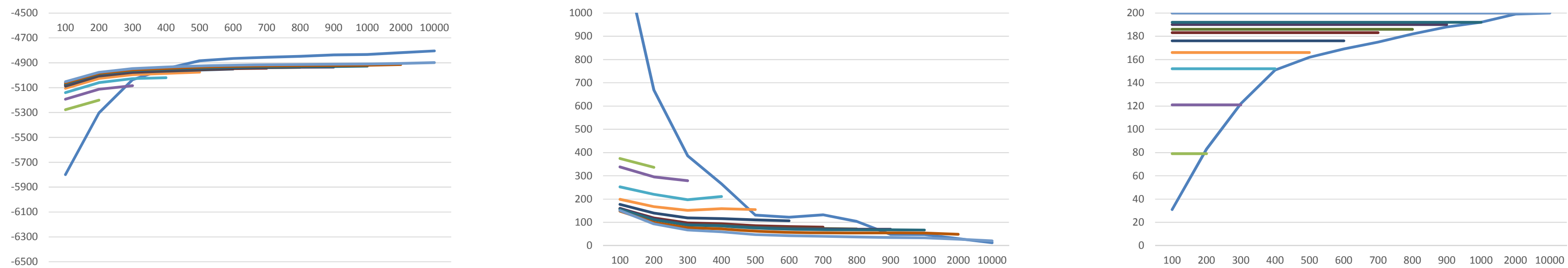
(a-1) L の平均　(b-1) L の標準偏差　(c-1) 残った b の回数

(1) y_1=1971 と y_2=1981



(a-2) L の平均　(b-2) L の標準偏差　(c-2) 残った b の回数

(2) y_1=1971 と y_2=1991

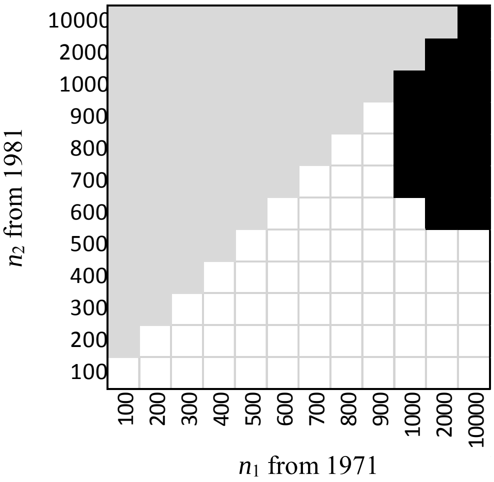


(a-3) L の平均　(b-3) L の標準偏差　(c-3) 残った b の回数

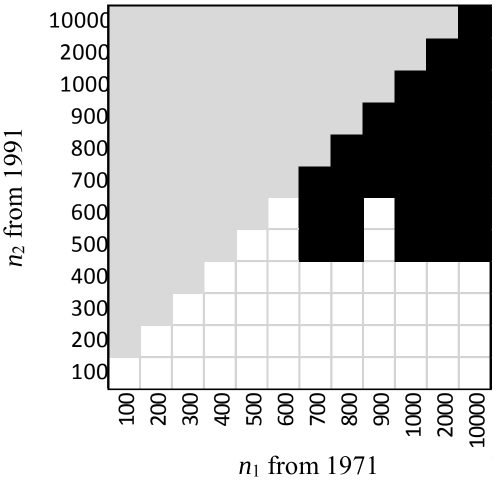
(3) y_1=1981 と y_2=1991

注：横軸は n_2．$L2$ と $L1(y_1, y_2, n_1=10000, n_2, b)$ の 2 つが 100〜10000 全体に描かれている線．どちらも青系統の色で描かれているが，濃いほうの図の中での触れ幅の大きいほうが $L2$.

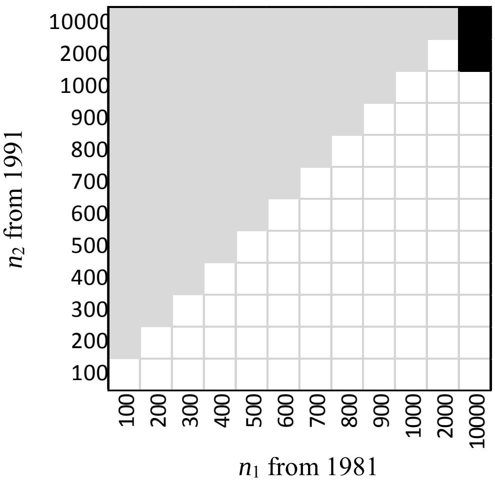
図—1　予測精度の特徴



(a) y_1=1971 と y_2=1981



(b) y_1=1971 と y_2=1991



(c) y_1=1981 と y_2=1991

注：灰色は n_1<n_2 であり，今回の興味の対象外．黒塗りは新しい時点のサンプルのみを用いたほうが定数項の修正よりも予測が有意に良い．

図—2　予測精度の差の検定

(a-1)～(a-3)のパネルを見ると，右肩上がりになっており，$L2$の場合にはn_2が大きいほど，また$L1$の場合にはn_1が同じであればn_2が大きいほど予測精度が平均して高いことが分かる．ところが$L2$の線が$L1$の線よりも上に現れることがある．これは，新しいデータのみを用いてモデルを構築したほうが古いデータを新しいデータで更新するよりも良い予測を平均して行えることを示している．

次に，$L1$と$L2$の標準偏差を図—1のパネル(b-1)～(b-3)に描く．今度は，右肩下がりであり，$L2$の場合にはn_2が大きいほど，また$L1$の場合にはn_1が同じであればn_2が大きいほど予測精度のばらつきが小さいことを示している．先ほどとは少し違うのが，$L1$と$L2$の線は平均値の場合よりも右で交差することである．つまり，n_2が小さいときには定数項の修正を用いたほうが予測精度は平均して高くそしてそのばらつきも小さい．n_2が増えると定数項の修正を用いたほうが予測精度は平均して低いがそのばらつきは小さい．さらにn_2が増えると定数項の修正を用いたほうが予測精度は平均して低くそのばらつきも大きくなる．

図—1のパネル(c-1)～(c-3)は，それぞれの場合で試みた200回のブートストラップのうち，何回で推定と予測が正しく行えたかを示したものである．パネル(c-3)において若干の例外が見られるものの，ほとんどの場合で定数項の修正を行う場合のほうが推定と予測を正しく行えることが多かった．

次に，データ年の違いによる考察をする．パネル(a-1)と(a-2)ではn_2=200程度以上において$L2$のほうが平均して高い予測精度をもたらすのに対して，パネル(a-3)ではn_2=500程度以上において$L2$のほうが平均して高い予測精度をもたらす．このことは，表—1において，修正前の1971年のモデルよりも修正前の1981年のモデルのほうが2001年の予測精度が極めて高いのに対し，修正前の1991年のモデルは修正前の1981年のモデルに比べて2001年の予測精度がそれほど大きくは向上していないことから解釈できる．つまり，古い時点のモデルの移転性が新しい時点のモデルの移転性に比べてそれほど低くなければ，2つの時点のデータを用いてサンプル数が増えることの効果が，古い時点のモデルの移転性の低い定数項以外のパラメータの悪影響を凌駕していると考えられる．

(3)　予測精度の差の検定

古い時点の多数のサンプルと新しい時点の少数のサンプルを両方用いて定数項を修正したモデルと，新しい時点の少数のサンプルのみを用いたモデル，の2つのどちらが予測精度がよいのかを検定した結果を図—2に示す．なお，ここで用いている統計は，繰り返しのbのうち，古い時点の多数のサンプルと新しい時点の少数のサンプ

ルを両方用いて定数項を修正したモデルと，新しい時点の少数のサンプルのみを用いたモデル，の両方で正しくモデルが構築されて，予測が得られた場合のみに基づいている．先ほどの図—1 (c-1)～(c-3)で*L*1または*L*2それぞれについて残っている繰り返し回数について示したが，ここでの残っている繰り返し回数は2つの方法を同時に考慮しているので注意が必要である．

1971年と1981年の場合をパネル(a)に，1971年と1991年の場合をパネル(b)に，1981年と1991年の場合をパネル(c)に示す．横軸に古い時点のサンプル数，縦軸に新しい時点のサンプル数をとった．各パネルの左上方にある灰色に塗りつぶされている領域は$n_1<n_2$（古い時点のサンプル数よりも新しい時点のサンプル数のほうが大きい）であり，今回の興味の対象外である．黒色で塗りつぶされている領域は$z > 1.96$となり，新しい時点の小さいサンプルを用いたほうが，2つの時点のデータを用いて定数項の修正をするよりも有意に予測精度が良いことを示している．なお，今回$z < -1.96$となることは1回もなかった．つまり，古い時点のサンプルと新しい時点のサンプルを両方用いる定数項の更新のほうが，新しい時点のサンプルのみを用いるよりも有意に優れた予測を行う場合は全くなかった．

図—2より，古い時点のサンプル数が一定程度あり，新しい時点のサンプル数も一定程度あるときに，新しい時点のサンプルのみで有意に良い予測を行うことができた．新しい時点のサンプル数が多いときというのは，1時点のデータのみで良い予測を行うためには一定程度以上のサンプル数が必要であることを意味している．古い時点のサンプル数が少ないときに棄却されないことは，古い時点のサンプル数が少ないと予測のばらつきが大きいからではないかと推測される．

また，いずれの時点においても古い時点のサンプル数がいくら大きくなっても，新しい時点のデータのみを用いた場合のほうが良い予測を得るために必要な新しい時点のサンプル数はほぼ一定であった（パネル(a)ではn_2=600～700程度，パネル(b)ではn_2=500～700程度）．このことは，新しい時点でのサンプル数をいくらにしたらよいか，という定量的な知見をもたらすのに役に立つと考えられる．

6. おわりに

統計的な検定によると，古い時点のサンプル数と新しい時点のサンプル数のどの組み合わせにおいても，定数項の修正による方法では新しい時点の少ないサンプルのみを用いるよりも良い予測を行うことができなかった．一方，古い時点のサンプル数と新しい時点のサンプル数が一定以上あるとき，新しい時点のサンプルのみを用いたほうが，定数項を修正するよりも統計的に良い予測を行うことができた．つまり，定数項の修正によるメリットは，定数項を修正することによって新しい時点のデータのみを用いるよりも統計的に良い予測を行うことができる，ということには求められない．

定数項の修正のメリットは以下の点にあると考える．

- 新しい時点のサンプル数が極めて少ない場合，定数項の修正のほうが良い予測を平均して行える．
- 新しい時点のサンプル数が少ない場合，定数項の修正のほうが予測のばらつきが小さい．
- ほとんどすべての場合において定数項の修正のほうが推定に関する問題が少ない．

なお，今回の事例は3つの選択肢からなる交通手段選択モデルというシンプルなモデル構造を持っていた．モデルの複雑性が与える影響についても検討する必要がある．

謝辞：本研究はJSPS科研費25380564の助成を受けている．データ使用に関して，中京都市圏総合都市交通計画協議会と名古屋大学森川研究室の支援を受けた．

参考文献

1) Dissanayake, D., Kurauchi, S., Morikawa, T. and Ohashi, S.: Inter-regional and inter-temporal analysis of travel behaviour for Asian metropolitan cities: Case studies of Bangkok, Kuala Lumpur, Manila, and Nagoya, *Transport Policy*, Vol. 19, No. 1, pp. 36–46, 2012.
2) Duffus, L.N., Alfa, A.S. and Soliman, A.H.: The reliability of using the gravity model for forecasting trip distribution, *Transportation*, Vol. 14, No. 3, pp. 175–192, 1987.
3) Elmi, A.M., Badoe, D.A. and Miller, E.J.: Transferability analysis of work-trip-distribution models, *Transportation Research Record*, 1676, pp. 169–176, 1999.
4) Sanko, N.: Travel demand forecasts improved by using cross-sectional data from multiple time points, *Transportation*, Vol. 41, No. 4, pp. 673–695, 2014.
5) Sanko, N., Dissanayake D., Kurauchi, S., Maesoba, H., Yamamoto, T. and Morikawa, T.: Inter-temporal analysis of household car and motorcycle ownership behaviors - The case in the Nagoya Metropolitan Area of Japan, 1981–2001 -, *IATSS Research*, Vol. 33, No. 2, pp. 39–53, 2009.
6) 三古展弘：交通需要予測におけるデータの新しさとサンプル数のトレードオフ，土木計画学研究・講演集，No. 50（CD-ROM），2014.
7) Sanko, N.: Trade-off between data newness and number of observations for travel demand forecasting, *Compendium of Papers of the 94th Annual Meeting of the Transportation Research Board*, Washington D.C., U.S.A., Jan. 2015.
8) Badoe, D.A. and Miller, E.J.: Comparison of alternative methods for updating disaggregate logit mode choice models, *Transportation Research Record*, 1493, pp. 90–100,

1995.
9) Karasmaa, N. and Pursula, M.: Empirical studies of transferability of Helsinki metropolitan area travel forecasting models, *Transportation Research Record*, 1607, pp. 38–44, 1997.
10) Atherton T.J. and Ben-Akiva, M.E.: Transferability and updating of disaggregate travel demand models, *Transportation Research Record*, 610, pp. 12–18, 1976.
11) Sanko, N. and Morikawa, T.: Temporal transferability of updated alternative-specific constants in disaggregate mode choice models, *Transportation*, Vol. 37, No. 2, pp. 203–219, 2010.
12) Efron, B. and Tibshirani, R.J.: An Introduction to the Bootstrap. Chapman & Hall, London, 1993.

（2015. 4. 24 受付）

SHOULD SMALL SAMPLES FROM RECENT TIME POINT BE USED WITH OLDER DATA?
APPLICABILITY OF UPDATING MODELS BY TRANSFER SCALING

Nobuhiro SANKO