

## 志願動向の隔年現象抽出の試み ——時系列データ解析の適用——

倉元直樹（東北大学高等教育開発推進センター）

志願者数の増減には各大学とも神経を尖らせ、志願者確保に多大な労力を注いでいる。一方、様々な攪乱要因によって志願動向が左右されている現実もある。本研究では 29 年間にわたる東北大学一般入試前期日程試験の志願者数等のデータに対して、時系列データ解析の手法を適用し、志願動向の隔年現象の析出を試みた。全受験者に対してはモデル適用の前提が満たされなかったが、特定学部に対しては時系列データ解析の手法が有効で、隔年現象が見出された。

### 1 問題と目的

#### 1.1 大学入学志願者数の問題について

高等教育のユニバーサル化が進む中、各大学とも、ありとあらゆる手段を尽くして志願者数の確保に努めるのが当然とされる風潮がある。そのため、入試広報活動は活発化の一途をたどり、放っておけば際限なく膨れ上がる傾向にある。しかし、広報活動には大学側のみならず高校側の負担感も大きい。本来は広報活動の費用対効果の分析と有効な広報に向けての戦略策定は欠かせない（倉元, 2007, 2008）。そういった認識は一部の関係者に共有されるようになってきた。そのため、最近ではいくつかの個別大学では広報活動の効果の評価に関する分析が試みられている（例えば、本多・島田・大谷・高野・関・佐藤・白川, 2011; 福島・吉村・坂本・笠原, 2011; 西郡・藤田, 2011）。

大学教育の質保証の前提として、いわゆるアドミッションポリシーに沿った学生の獲得は不可欠である。しかし「大学教育の質を担保するには、志願者の数を確保しなければならない」という信念が妥当なものか否かという点については疑問が残る。実際には、志願倍率に現れる表面的な人気と入学者の質が連動するものではないとする分析もある（植田・内海・平, 1996）。

実質的な議論を展開する以前に、そもそも受験生の志願動向を左右する要因は特定されていない。志願者数の増減や志願倍率の変化

が大学教育の本質と関係があり、広報活動の結果が数値に反映されるものならば、健全な状況と言えよう。しかし、大学入試関係者には、志願者数や志願倍率が教育の質や広報の努力等とはおよそ無関係な些末な要因によって大きく影響されているという経験則が根強く存在する。例えば、少子化に起因する大学進学適齢層の減少による影響は、避けようがない（西郡, 2012）。他にも受験生心理に影響して出願行動を左右する様々な攪乱要因の存在が信じられている（例えば、高木, 2012）。本来、議論の基盤として人為的努力が及ばない不可避的要因を排除した上で志願者数の増減を論ずるべきだろう。しかし、論拠となるデータや合理的に解析する手段が乏しいのが現状である。

そこで、本研究では志願者数等の経年データに対して時系列データ解析の手法を用いた分析を試みる。すなわち、時系列モデルの適用によって、通常状態における志願者数等の増減に関する法則的機序を見出し、議論の基盤を作ることを試みる。具体的には「志願者数や志願倍率に『隔年現象』は存在するのか」という問いに対して、時系列データ解析の手法を用いて検討を加えることを試みる。

#### 1.2 平成 24 (2012) 年度入試における東北大学への志願動向

研究課題の意義を説明するため、最初に、東北大学の入試制度について簡単に紹介する。

さらに東北大学の志願動向にかかわる要因として平成 24 (2012) 年度入試が置かれた状況に関する説明を加えることとする。

### 1.2.1. 東北大学の主要な学部入試の区分

東北大学では平成 2 (1990) 年度に工学部が推薦入学を開始した。平成 12 (2000) 年度には工学部と歯学部が AO 入試を開始した。しばらくは、推薦入学から AO 入試に切り替える学部、新規に AO 入試を導入する学部、推薦入学を導入する学部が混在していたが、徐々に AO 入試による入学者の入学後の成績が優秀である (倉元・大津, 2011) など、学力重視の東北大学型 AO 入試のメリットが各学部に浸透し、平成 19 (2007) 年度以降は一般入試後期日程や推薦入学を廃止して AO 入試を導入する方向性が定着した。平成 21 (2009) 年度の文学部を最後に、全学部で AO 入試が導入された。平成 24 (2012) 年度現在、全学 2,386 名の募集人員の内訳は、一般入試前期 1,860 名 (78.0%)<sup>1)</sup>、後期 93 名 (3.9%)、AO 入試Ⅱ期 158 名 (6.6%)、Ⅲ期 260 名 (10.9%)、推薦入試 15 名 (0.6%) である。

### 1.2.2. 東日本大震災の影響に対する懸念

平成 23 (2011) 年 3 月 11 日の東日本大震災は、平成 23 (2011) 年度一般入試後期日程の個別試験が予定されていた前日に発生した。東北大学では個別試験が中止となり、大学入試センター試験を主たる選抜資料として合否判定が行われた。結果的に入学辞退者が続出し、空前の辞退率を記録することとなった。東北大学の学部入試に対する、東日本大震災の目に見える直接的かつ最大の影響と言える。

平成 23 (2011) 年度に入り、入試広報活動の面でも間接的な震災の影響が懸念される出来事が続いた。震災後の混乱を治め、辛うじて例年通りに 7 月に大阪と東京で東北大学主催の受験生に向けた進学説明会が開催されたが、それまで右肩上がり順調に増加していた参加者数が一気に激減した。さらに、予定

通りに 7 月末に開催されたオープンキャンパスの参加者数も、前年の 51,766 名に対して 47,213 名と大幅に落ち込む結果となった。

### 1.2.3. 平成 24 (2012) 年度入試における東北大学への志願者数

平成 24 (2012) 年度入試における東北大学の志願者数の集計結果を表 1 に示す。全体を通して志願者数は前年度比 211 名減 (-2.7%) と「前年並み」あるいは「やや減少」と受け取られる程度の数値であり、事前に懸念されたほど大きな震災の影響が目に見える数値として表面化することはなかった。

表 1. 東北大学平成 24 (2012) 年度入試  
志願者数集計結果

	H23	H24	前年比	増減率
推薦入試	59 名	62 名	+3 名	+5.1%
AOⅡ期	412 名	427 名	+15 名	+3.6%
AOⅢ期	797 名	896 名	+99 名	+12.4%
一般前期	5,363 名	4,945 名	-418 名	-7.8%
一般後期	1,204 名	1,294 名	+90 名	+7.5%
合計	7,835 名	7,624 名	-211 名	-2.7%

入試区分別に見た場合、入試の種類で大きな違いが見出された。一般入試前期を除く入試区分では志願者数が増加した。特に AO 入試Ⅲ期では全学部の合計で史上最多の志願者数となった。一方、募集人員が最多の一般入試前期が前年度比 418 名減 (-7.8%) となり、全体での志願者減を導く原因となった。

通常の入試結果の検討過程では、以上のようなデータに学部別等にセグメント化した情報を加えて、何らかの意味を汲み取ろうとするのが一般的な手順であろう。数年間遡ったデータを参考として添付する場合もある。しかし、そのような表面的な数値を手掛かりに因果関係を類推してもっともらしい解釈を加え、結果を一面的に評価することだけで、果たして、的確な方向性を導くことが可能なの

だろうか。データの見方を変えれば、もしかすると同じデータから異なる様相が浮かび上がってくるかもしれない。

### 1.3 志願者数データの解釈における隔年現象

前年度の数値との比較に解釈を加えることは陥穽がある。志願動向に「隔年現象」が存在するならば、実質的な意味は見出せない。つまり、2年前と比較して前年度の数値が増加していたならば、今年度の数値が減少していても単なる隔年現象による通常の変動に過ぎないとも考えられる。逆もまた真なりと言えるであろう。

隔年現象ということばは、一般的に使われている学術用語ではなく造語である。本稿では「隔年現象」を長期的傾向の変化は除いて「1年おきに志願者数等が増減を繰り返す現象」と定義する。

隔年現象の存在を直近の数年間のデータから導くのは難しい。たとえ、数年連続して増加、減少が起こっているとしても、それが意味のある変化か否かは判断しがたい。より長期的な経年データを統計的に分析する必要があると思われる。

## 2 時系列モデルと隔年現象

### 2.1 自己回帰和分移動平均モデル (ARIMAモデル)

そこで、本研究では様々な分野で時間経過による変動の分析に用いられている時系列データ解析の手法（例えば、北川，2005）を志願者数等のデータに適用することを試みる。本稿で扱う経年データの解析に用いる手法は包括的には自己回帰和分移動平均 (ARIMA: Autoregressive Integral Moving-Average) モデル<sup>2)</sup> に属する。

ARMA (p, q) と表現される次数 (p, q) の自己回帰移動平均モデルは、(1)式で表される。

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (1)$$

ただし、 $t=1,2,\dots,T$ 、 $T$  はデータ系列の時点総数、 $\phi_i$  は  $i$  次の自己回帰母数 (AR 母数)、 $\varepsilon_j$  は時点  $j$  における平均 0、分散  $\sigma^2$  の攪乱項、 $\theta_j$  は  $j$  次の移動平均母数 (MA 母数) である。自己回帰母数を含む項と  $\varepsilon_t$  のみを含むモデルは自己回帰モデル (AR モデル)、移動平均母数を含む項と  $\varepsilon_t$  のみを含むモデルは移動平均モデル (MA モデル) と呼ばれる。いずれも  $p=0$ 、ないしは、 $q=0$  と置くことによって ARMA モデルに包摂される。

ARMA モデルを用いて時系列データの解析を行うには、攪乱項を含む確率過程が定常 (stationary) の条件を満たしていなければならない。しかし、志願者数や志願倍率などの入試データには、大域的 (global) に見て何らかのトレンドが見出されるのが普通である。また、長期的に同じ法則に従って変化が生じると考えるのも難しく、局所的 (local) にトレンドそのものが変化していく。

そのような場合、データの階差を取って局所的トレンドを除去し、時系列データを定常化する方法がある。

1 次の階差数列は (2)式で定義される。

$$\Delta x_n = x_n - x_{n-1} \quad (2)$$

これを  $d$  回くり返すと (3)式のように  $d$  次の階差数列ができる。

$$\Delta^d x_n = \Delta^{d-1} x_n - \Delta^{d-1} x_{n-1} \quad (3)$$

$d$  次の階差系列を作成し、ARMA (p, q) モデルを適用したものが ARIMA (p, d, q) モデルである。ARIMA モデルは、 $d=0$  と置くことで、ARMA モデルを包摂する。

### 2.2 ARIMA モデルにおける隔年現象

隔年現象とは、 $p=1$  の AR モデル AR (1) で自己回帰母数、すなわち、ラグ数  $\tau=1$  の自己相関係数  $\rho(1)$  ( $=\phi_1$ ) が負になるケースだと考えると分かりやすい。すなわち、平均  $\mu$  を 0 とした場合には、(1) 式を用いて

$$y_t = \phi_1 y_{t-1} + \varepsilon_t \quad (4)$$

とした場合に  $-1 < \phi_1 < 0$  となる 1 次の自己回帰母数を持つケースである。第 1 階差を用いた場合には ARIMA (1, 1, 0) と表される。

(9) 式のような  $q=1$  の MA (1) で移動平均母数が負になるケースも AR (1) と同様に解釈することができる。

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} \quad (5)$$

第 1 階差を用いた場合には、ARIMA (0, 1, 1) と表記される。

自己相関係数との関係は、Harvey (1981) / 国友・山本訳 (1985) によると

$$\rho(1) = \theta_1 / (1 + \theta_1^2) \quad (6)$$

となる。なお、より複雑な ARIMA モデルも理論的には可能だが、モデルの示唆する意味の解釈が難しくなることが難点だと思われる。

### 3 方法

#### 3.1 分析対象データ

昭和 59 (1984) ~平成 24 (2012) 年度の一般入試前期日程<sup>3)</sup>の大学全体ならびに A 学部の志願者数と志願倍率を用いることとした。

#### 3.2 分析手続き

ボックス＝ジェンキンスの方法 (例えば、Box and Jenkins, 1976) にしたがって、試行錯誤的にモデル識別 (identification)、推定と診断 (estimation and diagnostic checking)、予測 (forecasting) のプロセスを繰り返す。具体的にはおおむね以下のような手続きを経る。

- 1) 時系列データのグラフを作成し、トレンドがあるかどうかを視覚的に確認する。
- 2) コレログラムによる自己相関の変化を確認する。
- 3) 単位根検定、ホワイトノイズ検定を行い、階差をいくつまで取るべきか判断する。
- 4) 階差系列についてもトレンドの確認を行う。
- 5) 最適なモデルの候補を探す。
- 6) 適当な階差で複数の ARMA モデルの比較を行う。

- 7) 選択したモデルによる予測を行う<sup>4)</sup>。

分析手続きの 3) の段階は分析対象となる時系列データを ARIMA モデルの前提を満たすか否かを判断するために必要なプロセスである。

単位根検定では時系列データがトレンドを含む非定常なものか否かを判断する。帰無仮説が棄却されない場合には階差を取ったデータにこの手続きを繰り返す。ホワイトノイズ検定では時系列データが純粋にランダムな過程か時系列的な構造を有するかを判断する。帰無仮説が棄却された場合にモデル識別のプロセスに入る。

### 3.3 計算ソフト

主として SAS9.2 の ARIMA プロシジャを用いて分析を行った<sup>5)</sup>。

## 4 結果

#### 4.1 全学データの分析

全学データについては、志願者数、志願倍率いずれについてもフィリップス＝ペロンによる単位根検定 (Phillips-Perron test) の結果、階差を取る必要が見出された。そこで、第 1 階差を取ったが、6 次までの自己相関係数に対するホワイトノイズ検定の結果が志願者数 ( $\chi^2[6] = 6.11, p > .05$ )、志願倍率 ( $\chi^2[6] = 5.55, p > .05$ ) のいずれにおいても 5% 水準で有意とならなかった。そのため、その後の分析を断念することとした。

時系列モデル不成立の主因は、時点数の不足と思われる。今後、データの蓄積によって、分析に耐える水準に達することも期待できる。

#### 4.2 学部別データの分析

典型事例として、A 学部のデータに関して詳述する。図 1 は A 学部の志願倍率の変化である。明らかに複雑なトレンドと細かい年次変化の双方が見出される。

次いで 1 次階差系列を図 2 に示す。トレン

ドが消えて、隔年現象とみられる毎年の倍率変化が現れたことが視覚的に確認できる。

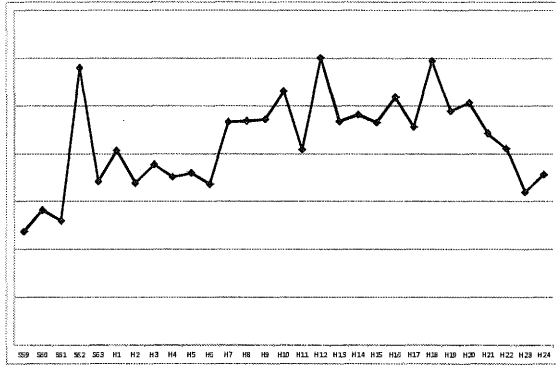


図 1. A 学部志願倍率の経年変化

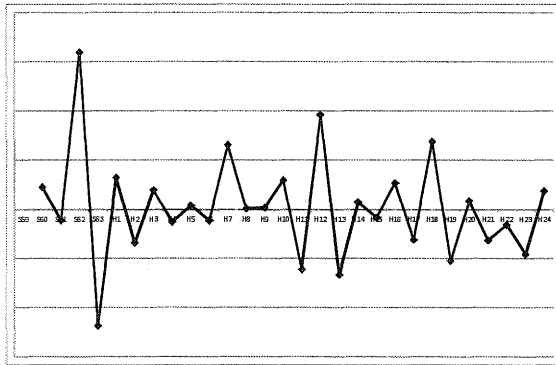


図 2. A 学部志願倍率の第 1 階差系列

図 3 にコレログラム (標本自己相関関数) を示す。ラグ数  $\tau=1$  の自己相関係数の値が負であり、 $\tau$  が大きくなるにつれて交互に正負に振動しながら減衰している。したがって、ARIMA (1, 1, 0) で  $-1 < \phi_1 < 0$ , または、ARIMA (0, 1, 1) で  $-1 < \theta_1 < 0$  となるモデルが適切であることが示唆される。

単位根検定, ホワイトノイズ検定 ( $\chi^2[6] = 22.43, p < .001$ ) の結果, 第 1 階差モデル ARIMA (p, 1, q) の妥当性が示唆された。次に, 適合する可能性があると思われるいくつかの第 1 階差モデルの適合度に対して AIC (Akaike's information criterion) 指標, BSC (Schwarz' Bayesian criterion) 指標で比較した。その結果を表 2 に示す。なお, 志願者数についても  $d=1$  のモデルが適切と判断され

た (ホワイトノイズ検定の結果は  $\chi^2[6] = 17.60, p < .001$ ) ので, その結果を同時に表 2 に示す。志願者数, 志願倍率の双方とも AIC, SBC の両指標において ARIMA (1, 1, 0) が最小となり, 最適モデルと判定された。

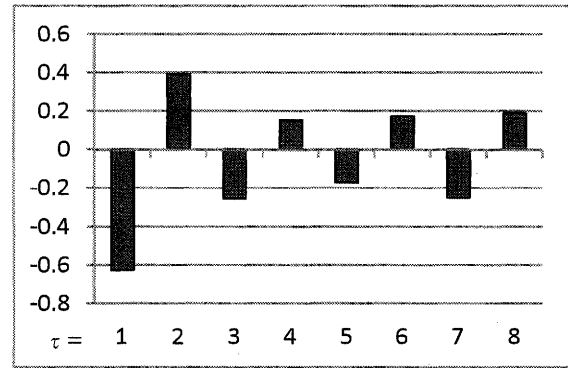


図 3. A 学部志願倍率 (第 1 階差) コレログラム

志願者数においては  $\hat{\phi}_1 = -.56$ , 志願倍率においては  $\hat{\phi}_1 = -.63$  となり, いずれも隔年現象の存在が実証された。なお, 残差のコレログラムからはホワイトノイズ過程が示唆されるプロットが得られた。

表 2. 時系列モデルの適合度 (A 学部)

モデル	志願者数		志願倍率	
	AIC	SBC	AIC	SBC
<b>ARIMA(1,1,0)</b>	<b>338.31</b>	<b>340.97</b>	<b>72.58</b>	<b>75.24</b>
ARIMA(2,1,0)	340.30	344.30	74.58	78.57
ARIMA(0,1,1)	338.57	341.24	75.95	78.61
ARIMA(0,1,2)	339.75	343.75	75.93	79.93
ARIMA(1,1,1)	340.30	344.29	74.58	78.57

## 5 考察

時系列モデルには, 長い期間の経年データの存在が必要とされる。各区分の入試は年に 1 回であるため, 時系列データ解析の手法が適用できるほど長い年月にわたるデータの入手は容易ではない。本研究で分析した一般入

試前期試験データもたかだか29年間であり、通常は分析に耐え得ないとされる長さであった<sup>9)</sup>。しかも、平成2(1990)年度には募集人員の分割による入試機会の複数化が断行されている。一般入試の機会が一度であった時期のデータを本質的に前期日程試験とみなして分析に組み入れ、辛うじて時系列モデルによる分析を成立させたという事情がある。つまり、本研究のデータの条件は非常に厳しいものだった。

それにもかかわらず、A学部では明確な隔年現象を析出された。このことから、場合によっては時系列モデルが入試データの解析に大きな威力を発揮する手法である可能性が示されたと言えるだろう。

先述のように隔年現象の存在は短いスパンのデータの解釈を誤らせる可能性がある。トレンドを除去した上で隔年現象を考慮した場合、大きな変化と見えたものが通常の揺らぎの範囲だったり、逆の意味を持つ可能性もあるからだ。本研究の分析では、A学部のデータにくっきりとした隔年現象が現れた。A学部の志願者は全国の大学の中から自分の学力に合わせた志望校を選ぶ傾向があると言われている。前年度の倍率や志願者数の数値が志望校選択の重要な情報源となっていると考えれば、隔年現象の存在は不思議ではない。

最後に本研究の分析の限界に触れておきたい。時系列モデルによる志願動向の分析は強力だが、あくまでも一つのツールである。一人ひとりの志願者にはそれぞれの状況があり、個別の事情に応じて進学の意味決定がなされることを忘れてはならない。大学入試の当事者として、入試データに対する科学的なアプローチを重視すると同時に一人ひとりの受験生を大切に見る姿勢を持つ続ける必要がある。

本研究の分析はまだ粗いものと言わざるを得ない。データをより精緻化して多変量モデルを適用するなどの工夫により、今回は見出せなかった時系列的構造が見出される可能性もある。いずれにせよ、本研究で用いた

方法を入試データの解析手法の一つの参考として活用していただけるならば、幸いである。

## 謝辞

時系列モデルの基礎に関しては、東北大学入試センター前センター長安藤朝夫先生(東北大学大学院情報科学研究科教授)にご指導いただきました。心から感謝いたします。

## 注

- 1) 医学部医学科の時限付き臨時増定員を含む。
- 2) 本稿の表記と用語は基本的に Harvey (1981) / 国友・山本訳 (1985) にしたがう。
- 3) 平成元(1989)年度までは前期と後期の区別が設けられていなかったため、一般入試全体のデータを用いることとした。
- 4) 本研究では、予測の段階には踏み込んでいない。
- 5) 本研究の計算には東北大学サイバーサイエンスセンター大規模計算システムを利用した。
- 6) SASのARIMAプロシージャのマニュアルには、「最低30以上のオブザベーションが必要」と記載されている。本研究の標本の大きさ  $T=29$  はどう考えても許容範囲ぎりぎりのサイズである。また、杉原(1986)は小標本でのARMA(1,1)過程においては $\phi$ と $\theta$ を正確に推定することが困難であることを指摘している。したがって、意味のある分析を行うためには、長いスパンのデータの蓄積と同時に可能な限りシンプルなモデルの適用が不可欠だと言える。

## 文献

- Box, G. E. P. & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*, revised edition, Holden-Day, San Francisco.
- 福島真司・吉村修・坂本嵩幸・笠原龍司 (2011). 「大学入試広報における効果測定の研究——データベースを用いた入試広報媒体の測定について——」『大学入試研究ジャーナル』21, 75-82.
- Harvey, A. C. (1981). *Time Series Model*, Philip Allan Publishers Limited, Oxford (国友直人・山本拓訳 [1985]. 『時系列モデル入門』東京大

学出版会).

- 本多正尚・島田康行・大谷奨・高野雄二・関三男・佐藤真紀・白川友紀 (2011).「大学の入試広報と入学者の利用する情報源の差異およびその評価」『大学入試研究ジャーナル』**21**, 69-74.
- 北川源四郎 (2005).「時系列解析入門」, 岩波書店.
- 倉元直樹 (2007).「東北大学入試広報戦略のための基礎研究 (1) —過去 10 年の志願者数・合格者数等から描く『日本地図』—」『東北大学高等教育開発推進センター紀要』**2**, 9-22.
- 倉元直樹 (2008).「東北大学入試広報戦略のための基礎研究 (2) —過去 11 年の志願動向に見る各募集単位の特徴—」『東北大学高等教育開発推進センター紀要』**3**, 63-76.
- 倉元直樹・大津起夫 (2011).「追跡調査に基づく東北大学 AO 入試の評価」『大学入試研究ジャーナル』**21**, 39-48.
- 西郡大 (2012).「18 歳人口の減少がもたらす入試の現状をどう把握するか? —今後の入試戦略に向けた基礎分析—」『全国大学入学者選抜研究連絡協議会第 7 回大会研究発表予稿集』(印刷中).
- 西郡大・藤田修二 (2011).「入試広報および入試改善に向けた情報収集—高校訪問活動から得られた知見—」『全国大学入試研究ジャーナル』**21**, 219-224.
- 杉原左右一 (1986).「ARMA 誤差を伴う回帰モデルの小標本特性—最尤推定量のバイアス評価—」『商学論究』**33** (3), 関西学院大学, 13-28.
- 高木繁 (2012).「センターリサーチと個別試験受験者の成績分布から見た『輪切り』の実態」『平成 24 年度全国大学入学者選抜研究連絡協議会第 7 回大会研究発表予稿集 (取扱注意)』, 独立行政法人大学入試センター, 69-74.
- 植田規史・内海爽・平直樹 (1996).「愛媛大学医学部における小論文入試への取り組みとその成果について」『大学入試センター研究紀要』**25**, 1-40.