大学入学希望者学力評価テストの複数回実施による合格者決定方法

濱口祐実　○村田忠彦（関西大学総合情報学部）

How to Grade Applicants by New University Entrance Examinations with Multiple Times

Y. Hamaguchi and * T. Murata (Department of Informatics, Kansai University)

概要－　本論文では，大学入学希望者学力テストにおいて，複数回実施を行った場合，受験生がもつ複数の点数を考慮した合格判定方法について，マイクロシミュレーションの技法を使って検討する．複数回受験において，受験生としては，自身の最良の得点を考慮した判定を希望する．一方，合否を判定する大学としては，受験生の実力を正しく判定する合格判定になることを期待している．本論文では，複数回の受験機会における受験生の真の学力の変化について，学習効果における変動と試験時の点数における変動の２つの変動を考慮してシミュレーションを行う．また，受験生の中に，学習効果が最終回の受験機会にだけ認められるものも考慮した上で，受験生の真の学力を判定する方法について検討する．数値実験により，複数回受験において，平均点，最高点，最低点を合格判定に用いる場合と１回受験による合格判定を比較したところ，１回受験により，十分，真の学力による合格者を判定できることがわかった．また，受験者数が少なくなり，相対的に合格率が上がった場合には，１回受験による合格判定の精度はより高くなることがわかった．

キーワード： 大学入学希望者学力テスト，マイクロシミュレーション，複数回受験，学習効果，試験時変動．

1　はじめに

計算科学は，理論科学，実験科学と並ぶ第３の科学として認められるようになっている [1]．物理学や化学，天文学などの分野で計算科学は利用されることが多いが，近年では，経済学や政治学，社会学の分野においても，社会シミュレーションという形で活用されるようになっている．人間をモデルに内包する社会シミュレーションは，物理学や化学などと比較して，不確定要素が多く，モデルの妥当性を確認する上で課題が残る場合が少なくないが，多様なシミュレーションを行うことにより，極端な事例（シナリオ）を発見し，それらの極端な事例から，避けなければならないシナリオや，望ましいシナリオを抽出し，それらのシナリオが発生した理由を解明し，そのメカニズムをモデルの範囲内で調べることにより，実世界で起こりうる現象の説明を行うとともに，避けなければいけない事象における予防，望ましい状況を生じさせるための施策の考案が期待される．

本論文では，2020年度から実施が予定されている「大学入学希望者学力評価テスト（仮称）」において複数回の受験機会が設けられた場合の学力評価方法に関して，マイクロシミュレーションの技法を用いて考察する．複数回受験は，体調不良や交通機関のトラブルなどで不利な結果となる場合がある１回受験と比較して，受験生にとって心理的な負担を軽減することが期待される．一方，受験生を受け入れる大学側にとっては，受験機会が複数になることにより，実施コスト，採点コストの増大が懸念される一方で，受験生の真の学力をより正しく判定するための方法としても期待されている．

大学入試の合格判定に関する理論的考察としては，河野[2]によるゲーム理論と確率モデルを用いた試論や，後藤ら[3]による確率分布モデルによる考察などがある．河野は確率モデルにより，受験者が２名の場合に，２回受験機会を与えたとしても，高々6%ほどしか，真の学力が上位のものを正しく判定する確率が上がらないことを示している．一方，後藤らは，合格と判定すべき受験生の得点分布の標準偏差と不合格と判定すべき受験生の得点分布の標準偏差が大きい場合には，合格者を不合格にしてしまう誤り確率が高くなるため，複数回の実施を行うとしても，合格者と不合格者を判別するための精度の高い試験の設計が重要になることを指摘している．

本論文では，複数回の試験実施に関して，マイクロシミュレーションによるアプローチを行う．河野や後藤らは，確率モデルや確率分布モデルを用いて，複数回試験の効果について検討を行っていたが，複数回受験の間に，受験生が学習を行うことによる真の学力の変化については検討していなかった．本論文では，学習による真の学力の変化を導入し，最終的な真の学力の判定を行うにあたり，複数回試験導入の効果に関する考察を行う．本論文では，受験機会における受験生の真の学力の変化を，一様乱数や正規分布で表現し，学力変化の状況に応じた適切な合格者判定方法について検討する．

2　複数回の受験結果の活用方法

複数回受験を認めている例として，アメリカ合衆国では，SAT（元は Scholastic Assessment Test の略称だが，今は SAT として呼称している）の複数回受験を認めている．2009 年以来，Score Choice という制度により，受験生が複数回試験の最もよい結果を選んで志望する大学に送ることができるようになっているが，SAT を利用している大学の多くが科目ごとに最高点を採用することを表明している [4]．SAT に関しては複数回受験が認められ，その成績についても，受験生にとって利用しやすいように配慮されているが，テストの成績以外にも，小論文や課外活動，面接の比重も大きいため，日本における大学入試制度との単純な比較は難しい．

複数回にわたって実施された試験結果を各大学がどのように活用するかは，合否を判定する大学に委ねられている．複数回受験の結果の活用方法には，最高点以外にも，平均点，最低点を活用する方法も考えられ

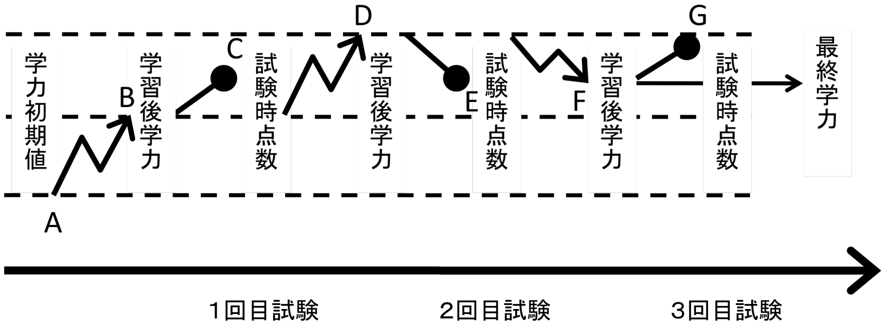

Fig.1：学力変動と試験時変動

る．受験生の合否を判定する大学としては，受験生の学力を正しく判定し，優秀な学生を合格にすることが目的である．本論文では，複数回受験の結果の活用方法による，合格者の判定への影響をマイクロシミュレーションの技法を用いて確認する．

3　モデル概要とシミュレーション結果

本研究では，100人の受験生がn回受験をし，上位20位以内の受験生が合格する，という状況を検討する．n回の受験機会の前にはそれぞれ学習機会が設けられ，学力が変動するものとする．

Fig.1に試験回数 $n = 3$ の場合の学力と試験時点数の変動の例を示す．まず，受験生各自は，学力初期値を持つものとする．今，ある受験生の真の学力を計測できるものとし，その点数がA点であるとする．学習することにより,その受験生の真の学力がB点に変動したものとする．１回目の試験時には，受験生の学力が正しく計測できないこともあるため，試験時の点数の変動として，C点が計測できたとする．２回目の学習の開始時には，受験生の真の学力はB点であるので，B点からの変動として学習後の真の学力がD点になる．以下同様にして，２回目の試験時の点数がE点，３回目の学習後の真の学力がF点，３回目の試験時の点数がG点で与えられるとする．このとき，受験生の合否を判定する大学としては,最終的な学力としてのF点をもとに合否を判定したいものとする．一方，試験により計測できる点数は，C点，E点，G点の３つの点数であるので，大学が行いたいことは，このC, E, Gの３つの値から，正しくFを推定することである．

なお，本研究における真の学力は，受験生間で相対的に比較可能なものであるとし，一定範囲内に正規化されたスコアであるとする．したがって，Fig.1の例の受験生は，３回目の学習後に学力がDからFに変化しているが，これは，真の学力が絶対値として低下した，というよりも，他の受験生との相対的な値として学力が低下した，と捉えることとする．

3.1 一様乱数を用いた予備実験

まず，各受験生に[10,100]の初期学力を一様乱数により与えて，試験回数を $n = 5$ としてシミュレーションを行った．以下に，まず，受験生の学習機会と学習時の変動,試験時の変動幅の値を次のように設定した．

・受験生：以下の２通りの学習傾向をもつ受験生群XとYを設定する．

Table 1：学習効果と試験時変動幅
（一様分布，受験者100人，合格者20人）

ケース	タイプ	学習効果	試験時変動幅
A	X 100人	[-10,10]	[-5,5]
B	X 100人	[-5,5]	[-5,5]
C	X 70人	[-10,10]	[-5,5]
	Y 30人	[-10,30]	[-5,5]
D	X 70人	[-10,10]	[-20,20]
	Y 30人	[-10,30]	[-20,20]

[X] 全ての学習機会で学習を行う

[Y] 最後の試験の直前のみ学習する

・学習効果：第t回目の学習後の真の学力 A_t （t=1, 2,. . , n）を次のように定義する．

$$A_t = A_{t-1} + [a,b] \tag{1}$$

ただし，$[a,b]$は下限a，上限bからの一様乱数の値．

・試験時変動（幅）：第t回目の試験時の点数S_tを次のように定義する．

$$S_t = A_t + [c,d] \tag{2}$$

ただし，$[c,d]$は下限c，上限dからの一様乱数の値であるとする．

次に，下記のA～Dの4つのケースを設定した．学習効果として$[a,b]$の値を，試験時の変動幅として$[c,d]$の値をTable 1のように示す．Table 1の4つのケースの設定により，全員が定期的に学習する場合（A, B）と，課外活動などに専念し，学習機会のうち最後の機会のみに学習する受験生がいる場合（C, D）との比較や，真の学力と試験結果の変動幅が小さい場合(A,B,C)と，受験生の体調が悪くなる場合や，試験の難易度などによって真の学力と試験結果の変動幅が大きくなる場合（D）との比較を行う．

受験生の合否を判定する大学の目的は，最終的な真の学力上位の受験生を合格にすることである．なお，「最終的な真の学力」とは，最終回の試験直前の学習機会を終えた後の受験生の真の学力である(Fig.1の例ではF点）．そこで，試験結果をもとに，受験生を順位付けした場合に，合格者の中に真の学力で獲得すべき

Table 2：試験の合格者と真の学力上位者の一致率
（一様分布，受験者100人，合格者20人）

	1回	平均点	最高点	最低点
A	92.91	82.57	84.47	84.50
B	91.92	88.40	89.45	89.32
C	92.54	78.70	84.72	80.19
D	75.68	75.83	80.46	75.82

Table 3：学習効果と試験時変動幅（正規分布）
（正規分布，受験者100人，合格者20人）

ケース	タイプ	学習効果	試験時変動幅
A’	X 100人	平均0，SD10	平均0，SD5
B’	X 100人	平均0，SD5	平均0，SD5
C’	X 70人	平均0，SD10	平均0，SD5
	Y 30人	平均10，SD20	平均0，SD5
D’	X 70人	平均0，SD10	平均0，SD20
	Y 30人	平均10，SD20	平均0，SD20

SD：標準偏差

Table 4：試験の合格者と真の学力上位者の一致率
（正規分布，受験者100人，合格者20人）

	1回	平均点	最高点	最低点
A’	90.58	76.03	80.04	73.70
B’	89.26	82.15	85.30	84.79
C’	91.04	70.31	82.63	68.65
D’	68.95	65.55	67.74	61.13

受験生を含む割合を，次のように一致率として定義する．

一致率＝試験合格者に含まれる真の学力上位者の数÷真の学力上位者の数　(3)

今，入学試験の合格者数を20人とする場合，試験結果で上位20位の受験生（合格者）の中に，真の学力で上位20位の受験生が18名含まれている場合，一致率は90%となる．したがって，一致率が100%になれば，真の学力上位者が全員，試験による合格者に含まれていることになる．

今，大学が計測できるのは受験生の試験時の点数だけ（Fig.1の３回試験の例では，C点，E点，G点）であるので，複数回試験の平均点，最高点，最低点に基づいて，合格者を判定する場合の一致率の違いをシミュレーションにより確認する．また，比較のため，最終試験１回だけの試験により合格者を判定する場合の一致率についても求める．

Table 2に各ケースでの最終試験1回（第5回目の試験結果），5回の試験の平均点，最高点，最低点に基づく一致率の100試行平均値を示す．

Table 1から，４つのケースのうち３つのケースA，B，Cで最終試験１回の順位による合格者決定で，最も真の学力との一致率が高いことがわかる．また，これらのケースでは，平均点による判定が最も低い一致率となっていることがわかる．これは，複数回受験では，各回の学習後の真の学力と最終回の真の学力がずれていることに加えて，試験結果は直前の真の学力に対して変動が加わる値になり，各回の試験結果に二つのズレが蓄積されるため，平均による合格者が，最終的な真の学力に基づく上位者と異なるものになると考えられる．これにより，試験結果の変動幅が小さい状況下では，複数回受験をしなくとも，1回試験で受験生の最終的な真の学力を測定できることがわかる．

また，ケースAとケースCの比較から，最終試験の直前にしか学習しないグループがいる場合も，試験結果の変動幅が小さいため，最終試験１回のみで，最も高い一致率を実現できることがわかる．

しかし，ケースCとケースDの比較からわかるように，真の学力と試験結果の変動幅が大きくなる場合，1回試験の一致率は悪くなる．これは，1回しか受験機会がないため，真の学力から大幅にずれた試験結果がでることにより，真の学力の測定が難しいためである．このような状況下では，複数回受験の最高得点を活用する方法が，最も正しく最終的な真の学力を測定できることがわかる．

3.2 正規分布を用いた数値実験

前節では，予備実験として，受験生の学力，学習効果，試験時変動幅において，一様乱数を仮定して受験生の得点の分布を与えた．一般的には，受験生の学力分布は大数の法則にしたがって，正規分布になっていると考えられており，河合塾の大学入試センター試験概況分析[5]においても，平均点をピークとする正規分布様の得点分布が得られていることが示されている．そこで，本節では，受験生の初期学力を平均50，標準偏差20の正規分布により与える．

次に，一様分布の場合と同様に，Table 3のように４つのケースを想定した．学習効果と試験時の変動幅についても，正規分布による平均値と標準偏差（表中のSD）により変動を与えている．Table 1の一様乱数のときの条件と同様に，全員が定期的に学習する場合（ケースA’，ケースB’）と，課外活動などに専念し，最後の学習機会のみに学習する受験生がいる場合（ケースC’，ケースD’）との比較や，真の学力と試験結果の変動幅が小さい場合（ケースA’，ケースB’，ケースC’）と，受験生の体調が悪くなる場合や，試験の難易度などによって真の学力と試験結果の変動幅が大きくなる場合（ケースD’）との比較を行う．

Table 4に数値実験の結果を示す．Table 4の結果からわかるように，正規分布の場合，全てのケースで１回試験の一致率が最も高くなった．これは，一様乱数と比較して正規分布の方が，得点の変動が均一になるため，真の学力に対して，学力の変動が起こりにくく，最終試験１回の合格判定が最も一致率が高い結果になったと考えられる．

ケースA’とケースB’の比較により，正規分布を使った場合には，一様分布の場合と比較して，平均点による合否判定による一致率が最も悪い結果になっているわけではないことがわかる．１回試験の一致率が高い理由と同様に，これも正規分布により，学習効果や試験時の変動に関して，真の学力に近い点数が与えられる傾向が，一様乱数と比較して強くなるからであると考えられる．

Table 5：学習効果と試験時変動幅（正規分布）
（正規分布，受験者40人，合格者20人）

ケース	タイプ	学習効果	試験時変動幅
A''	X 40人	平均0, SD10	平均0, SD5
B''	X 40人	平均0, SD5	平均0, SD5
C''	X 28人	平均0, SD10	平均0, SD5
	Y 12人	平均10, SD20	平均0, SD5
D''	X 28人	平均0, SD10	平均0, SD20
	Y 12人	平均10, SD20	平均0, SD20

SD: 標準偏差

Table 6：試験の合格者と真の学力上位者の一致率
（正規分布，受験者40人，合格者20人）

	1回	平均点	最高点	最低点
A''	94.10	86.44	87.47	86.94
B''	93.71	88.88	91.33	90.98
C''	94.88	84.27	87.44	84.92
D''	81.00	80.71	80.19	78.50

ケースA'とケースC'の比較により，最終試験の時のみ学習する受験生がいる場合には，最終試験１回による合否判定がもっとも一致率を高めることがわかる．ケースD'のように試験時の変動幅が大きい場合には，どの判定方法を使っても一致率が７割を切る結果になってしまっているため，３割以上の本来合格にすべき受験者を不合格にしてしまっていることになる．

3.3 受験者数が減少する場合

前節まででは，受験者が100名，合格者が20名の場合の一致率を検討した．本節では，受験者数が少なくなる中で，合格者に変化がない場合の実験を行う．これは，18歳人口が，2010年の122万人から2030年の101万人に減るため，受験者数自体が減っていくことが想定される．そこで，入学定員は同じ間まで，受験者数が40名になった場合のシミュレーションを行う．

Table 5に，Table 3と同じ４つのケースを，総受験者数40名にして設定した．具体的には，ケースA''とケースB''は全ての受験者がXの場合で，ケースC''とケースD''では，Xが28人とYが12人になっている．それ以外の学習効果と試験時変動幅はTable 3と同じである．

Table 6の結果から受験者数が少なくなり，合格率が相対的に上がる場合において，全てのケースで，最終試験の１回の試験での合格判定により，最も高い一致率を与えることがわかる．Table 4とTable 6を比較することにより，受験者数が少なくなった方が，最終試験１回での一致率がより高くなっていることがわかる．

4 おわりに

本研究では，大学入学希望者学力評価テストにおいて，複数回実施が実現した場合に，各大学が受験生の学力を正しく判定するための受験結果の活用方法をシミュレーションにより考察した．具体的には，学習効果と試験時の変動を一様乱数と正規分布を用いて表現し，複数回受験の平均点，最高点，最低点，および最終試験１回での判定の４つの合格判定方法について比較を行った．また，受験生については，全員が全期間にわたって学習するケース，全員が全期間学習するものの，変動幅が小さいケース，７割の受験生が全期間学習し，３割の学生が最終試験の前にだけ学習する場合で，試験時の変動が異なるケースを２つ考慮した．これら４つのいずれのケースにおいても，一様乱数を用いた最終試験の前にのみ学習する受験生がいる場合で，試験時の変動が大きい場合のみ，最高点の判定法がもっともよかったものの，それ以外のケースでは，全て，最終試験１回の成績による合格判定がもっとも一致率を高める結果となった．

今回のシミュレーション結果から，複数回試験を導入したとしても，合格判定を行う大学が本来求めている真の学力上位者を合格にするための制度にはならない可能性が高いことを示している．受験生にとっては，複数回受験できることにより，試験時の緊張を和らげられることや，体調不良の場合の取り返しができる安心感があることなど，利点があるかもしれないが，試験を運営する大学側にとっては，合否判定にそれほど大きな影響がないことから，試験実施のコストの方がより負担になることが懸念される．この結論は，河野[2][河野02]や後藤ら[3][後藤06]の結論と同じである．

本研究での課題として，実データとしての得点分布が正規分布であることと，学習効果や試験時変動において，真の学力をもとに正規分布により変動を加えることが正しいモデルといえるかどうかの検討が必要である．学習モデルについては，矢野ら[6][矢野15]が提案している学力向上モデルを適用することも可能である．これには，真の学力と試験の成績との関係を明らかにする必要があり，真の学力の計測方法自体が課題となる．

また，今回は合否判定を１科目で行ったが，複数科目の場合の合否判定に与える影響についての検討も行うことができる．アメリカ合衆国で行われているSATでは，複数科目のそれぞれについて，複数受験の結果の最高点を使用できることが多く，最高点の組合せを用いることで，どの程度，それぞれの科目に対する真の学力の計測が可能になるかを検討する事も可能である．

参考文献

1) 小柳: 計算科学とシミュレーション，http://olab.is.s.u-tokyo.ac.jp/~oyanagi/reports/cs-and-sim.txt (2003年9月執筆)
2) 河野，大学入試制度における資格試験化と受験機会複数化の数理分析―ゲーム理論的視点からの試論―，理論と方法，Vol.17，No.2，195/209 (2002)
3) 後藤，石田，平澤，統計モデルに基づく大学入試の理論的考察，経営情報学会 2006 年秋季全国研究発表大会，224/227 (2006)
4) College Board, Score Choice - What if I decide not to use Score Choice? -, https://collegeread iness.collegeboard.org/sat/scores/sending-scores/score-choice (2016)
5) 河合塾，2016 年度大学入試センター試験概況分析，http:// www.keinet.ne.jp/topics/15/20160205-2.pdf
6) 矢野，神澤，山田，吉川，寺野：少人数学級は有効か？―エージェントベースシミュレーションによる学級定員と学力との関係，教育システム情報学会誌，Vol. 32, No. 4, 236/245 (2015)