統計モデルに基づく大学入試の理論的考察

A model of entrance examination of university based on statistical theory

後藤 正幸† 石田 崇‡ 平澤 茂一‡
Masayuki GOTO† Takashi Ishida‡ Shigeichi Hirasawa‡

†武蔵工業大学 環境情報学部
† Fac. of Environmental and Information Studies, Musashi Institute of Technology
‡早稲田大学 理工学部
‡ School of Science and Engineering, Waseda University

要旨:

大学市場の競争激化により、入学試験の位置づけが変わりつつある。大学入試は優秀な受験生を選抜するだけでなく、受験生へのメッセージとなるマーケティングツールである。本発表では、統計的モデルに基づき、入学試験の性質について理論的に分析を行い、その本質的な挙動を示す。これにより、大学入試において考慮に入れるべき事項を明らかにすると共に、将来的な入試のあり方についても言及を行う。

Abstract:

Recently, the market surrounding universities in Japan is changing and the competition to survive has intensified. The entrances examination is a tool of marketing of universities which give examinee messages. In this paper, we propose a statistical model of entrance examination. From the consideration based on the statistical model of the entrance examination, several suggestions can be given.

1. はじめに

近年、大学は少子化と大学定員数全体の増加に伴い、厳しい市場競争下に置かれようとしている。独立行政法人 大学入試センター[1]の情報によれば、センター試験の全国現役志願率をみても、毎年2万人程度の受験生減少が続いている。競争下におかれた大学では、市場に焦点を当てた大学マーケティングのパラダイムが必要とされることは自明とも言えるであろう[2]。一方、大学における顧客選定の方法としての入学試験は、大学マーケティングが有する特殊性の一つといえる。しかし、以前のように大学定員よりも志願者の方が遥かに多い状態では、入試は多くの大学においてその機能を果たしていたが、最近では受験生市場の縮小を受け、各大学は様々な差別化戦略を打ち出す大学が増加すると共に、大学入試のあり方も多種多様になりつつある。

本研究では、このような現在見られる入学試験の形態について、確率モデルや検定の概念を用いたモデル化を試み、大学入試の形態が与える選抜結果への影響について考察する。入学試験に対する数理的 統計的モデルを用いた議論としては、例えば、多属性効用分析の手法を用いた入学試験合格者数決定方法、大学入試データの解析方法論[3]などがある。また、試験そのものの設計法を対象とした理論には、項目反応理論に基づく議論がある。さらに、言語テストの問題作成と評価法、将来的な試験解答の自動分析法を論じた枠組みとして、言語モデルやテキストマイニング手法に基づく記述式試験解答の自動採点法などの議論もある。しかしながら、入学試験の方法をマクロ的にとらえ、入学試験の実施方法による影響を論じた研究はほとんどない。本稿では、入学試験の結果を確率モデルによって表現し、各受験生の試験結果は統計的データであるという認識に基づいた議論を行う。本稿で述べる結果は、大学マーケティングにおいて重要な入学試験に対して、基礎となる知見を与え、様々な入試戦略を立案するための基礎を与えるものである。

2. 大学市場を取り巻く現状と入試

大学経営にとって、少子化の影響が深刻な課題として認識されるようになった。全国では、実際に定員未充足、いわゆる定員割れの状況にある大学は 100 を越え 2001 年度調査時で 149 校)、現状では定員を満たしている大学も安泰とは言えなくなってきた。このような状況下、大学入試の形態は多種多様化している。例えば、一般的な大学の入学方法を考えてみても、指定校推薦、AO入試、大学入試センター利用試験、前期日程試験、後期日程試験、特待生入試、社会人入試、外国人入試など、様々な入試形態が存在する。センター利用試験、前期日程試験、後期日程試験といった形態を併用する大学が多くなり、事実上、志望する同一学部 学科にチャレンジする機会は増大した。その結果、従来見られた他大学 学部との併願数は減少し、受験生は志望度の高い大学 学部に絞って受験する傾向が強まったとの指摘もある。

例えば、明治大学では 2007 年度から全 8 学部共通の統一入学試験を実施し、一度の受験で複数の学部への出願を可能とする方法を実施すると発表している。複数の受験日から受験する日を選ぶことができる受験日自由選択制や同一志望学科を複数回受験できる同一学科複数日受験制といった方式もあり、例えば、東京農業大学ではⅠ期一般入試においてこの両方を採用し、受験生が 3 日間という試験日を活用できる入試を提供している。入試科目数に変化をもたせ、他科目にわたる総合的学力の高い学生だけでなく、少数教科に秀でた学生も採用しようという流れもある。例えば、愛知淑徳大学や大阪経済法科大学では一般入試の C 方式において、少数名の募集ながら1教科入試を実施している。2 教科型の入試も広がりを見せており、例えば法政大学の一般入試においても、3 科目入試 A 方式)と2 科目入試 B 方式)を実施し、試験日が異なれば、同一学部 学科を複数回受験できる方式となっている。これに対し、逆に従来の1教科型を廃止し、2 教科型に増やす大学もあり、各大学において入試教科に対する議論が続いているようである。

3. 受験者分布に関する検討

予備校などの受験業界では、大学入試に関する情報、とくに各大学学部の難易度を比較検討するために、受験者の模擬試験偏差値と実際の入試における合否判定を基にしたデータを分析し、各方面に提供している。ここでは、受験者の偏差値という指標を用いて、大学の難易度や受験生の学力を測る方法が一般に用いられる。これは実際の得点分布が正規分布に近い左右対称の単峰形であることを理由に有用な指標となり得ている。本章では、受験生の得点分布を正規分布と仮定した場合について議論し、得られるいくつかの知見について述べる。

3-1. 受験者全体の得点分布に関する検討

受験生全体の 1 回の試験結果に対する正規性について考えてみよう。ここで、試験の得点可能範囲が得点分布の分散よりも十分広い理想的な試験が行われることを仮定する。一人の受験生の得点が正規分布 $N(\mu_s, \sigma_s^2)$ に従うものとする。受験生の真の学力を μ_s と考え、各受験生の得点分布の分散を同じとみなせるとする。このとき、受験生の真の学力分布も正規分布 $N(\mu, \sigma^2)$ に従うものと考えれば、ベイズ統計学でよく知られる混合分布の計算式により、受験生全体の 1 回の試験結果に対する分布は、

$$p(x) = \frac{1}{\sqrt{2\pi(\sigma^2 + \sigma_S^2)}} \exp\left\{-\frac{(x-\mu)^2}{2(\sigma^2 + \sigma_S^2)}\right\} \quad (1)$$

となり、これは正規分布 $N(\mu, \sigma^2 + \sigma_S^2)$ で与えられる。受験生の学力が正規分布 $N(\mu, \sigma^2)$ に従えば、受験生の得点分布はその分散が $\sigma^2 + \sigma_S^2$ に増加した正規分布に従うことになる。すなわち、受験生の一発試験における得点誤差分だけ、得点分布はばらつきが大きくなる。

3-2. 合格すべき受験生群と不合格にすべき受験生群の得点分布に関する検討

つぎに、受験生の真の学力を $\theta = \mu_S$ とおき、その分布が $N(\mu, \sigma^2)$ に従うものとする。正規分布 $N(\mu, \sigma^2)$ に従う確率変数 θ の確率密度関数を $f_N(\theta)$ とし、正規分布 $N(\theta, \sigma_S^2)$ に従う確率変数 x の確率密度関数を $f(x|\theta)$ とする。いま、受験生全体の学力レベルにより、本来"合格すべき群"と"不合格にすべき群"の二群に分けることを考える。すなわち、あるしきい値 T を用いて、$\theta \geq T$ は合格すべき群、$\theta < T$ は不合格にすべき群と考えよう。このとき、本来、合格すべき群全体の確率密度関数 $p_1(x)$、及び不合格にすべき群全体の確率密度関数 $p_2(x)$ は、$f_N(\theta)$ の分布関数 $F_N(\theta)$ を用いて次式で与えられる。

$$f_1(\theta) = \begin{cases} f_N(\theta) \big/ F_N(T), & \theta \geq T \\ 0, & \theta < T \end{cases} \quad (2)$$

$$p_1(x) = \int_{-\infty}^{\infty} f_N(x|\theta) f_1(\theta) d\theta \quad (3)$$

$$f_2(\theta) = \begin{cases} f_N(\theta) \big/ (1 - F_N(T)), & \theta \leq T \\ 0, & \theta > T \end{cases} \quad (4)$$

$$p_2(x) = \int_{-\infty}^{\infty} f_N(x|\theta) f_2(\theta) d\theta \quad (5)$$

したがって、

$$p(x) = F_N(T) p_1(x) + (1 - F_N(T)) p_2(x) \quad (6)$$

$T \to -\infty$ のとき $p_1(x)$ が、$T \to \infty$ のときは $p_2(x)$ が正規分布 $N(\mu, \sigma^2 + \sigma_S^2)$ に近づく。すなわち、受験者のうちほとんどが合格になるケース ($T \to$ 小)、受験者のうちほとんどが不合格になるケース ($T \to$ 大) は、"合格すべき群"と"不合格にすべき群"の二群に分けたあとも、それぞれ正規分布と考えてよい状況となる。逆に、合格ラインがちょうど学生全体の学力の平均値に近づく場合、すなわち $T \to \mu$ の場合にどうなるかについて考えてみる。$T = \mu$ のとき、自明に

$$p(x) = \frac{1}{2}\{p_1(x) + p_2(x)\}$$

が成り立つ。さらに計算すれば、

$$p_1(x) = \frac{1}{\pi\sigma\sigma_S} \int_{\mu}^{\infty} \exp\left\{-\frac{(x-\theta)^2}{2\sigma_s^2} - \frac{(\theta-\mu)^2}{2\sigma^2}\right\} d\theta$$

$$= \frac{1}{\sqrt{2\pi(\sigma^2 + \sigma_S^2)}} \exp\left\{-\frac{(x-\mu)^2}{2(\sigma^2 + \sigma_S^2)}\right\}$$

$$\cdot \sqrt{\frac{2(\sigma^2 + \sigma_S^2)}{\pi\sigma^2\sigma_S^2}} \int_{\mu}^{\infty} \exp\left\{-\frac{\sigma^2 + \sigma_S^2}{2\sigma^2\sigma_S^2}\left(\theta - \frac{\sigma^2(x-\mu)}{\sigma^2 + \sigma_S^2}\right)^2\right\} d\theta$$

ここで、

$$\int_{\mu}^{\infty} \exp\left\{-\frac{\sigma^2 + \sigma_S^2}{2\sigma^2\sigma_S^2}\left(\theta - \frac{\sigma^2(x-\mu)}{\sigma^2 + \sigma_S^2}\right)^2\right\} d\theta = \sqrt{\frac{\pi\sigma^2\sigma_S^2}{2(\sigma^2 + \sigma_S^2)}}$$

であり、これは $T \to -\infty$ のときに、$p_1(x)$ が正規分布 $N(\mu, \sigma^2 + \sigma_S^2)$ に近づくことを示している。また、

$$\frac{\sigma^2}{\sigma^2 + \sigma_S^2}(x - \mu) \to 大$$

のとき、$p_1(x)$ が正規分布 $N(\mu, \sigma^2 + \sigma_S^2)$ で比較的精度よく近似できることもわかる。逆に、

$$\frac{\sigma^2}{\sigma^2 + \sigma_S^2}(x - \mu) \to 小$$

となると、正規分布 $N(\mu, \sigma^2 + \sigma_S^2)$ の分布の裾よりも急速に 0 に近づく分布形をしていることになる。(8)式より、$p_1(x)$ は、$\exp\{-(x-\mu)^2\}$ のオーダの関数 $O(\exp\{-(x-\mu)^2\})$ を用いて、

$$p_1(x) = \frac{1}{\sqrt{2\pi(\sigma^2 + \sigma_S^2)}} \exp\left\{-\frac{(x-\mu)^2}{2(\sigma^2 + \sigma_S^2)}\right\} \cdot O(\exp\{-(x-\mu)^2\})$$

という形で書き表せる。これはモード μ を持ち、左右非対称の単方形分布の形状をしていることがわかる。

3-3. 合格者分布と不合格者分布

予備校では毎年、在校生に対する受験結果と進学先の調査を実施しており、その集計結果から模擬試験の偏差値分布を知ることが可能である。これらの集計では、各大学の合格者と不合格者別で層別された偏差値分布なども提示され、大学ではその情報を得ることも可能である。その分布は、概ね下図に示すような形状となっている。

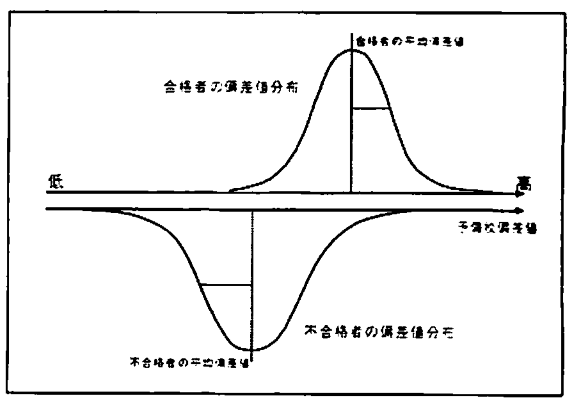

図 1．合格者と不合格者の偏差値分布

これは、合格者と不合格者の割合がほぼ等しい実質倍率が約 2倍程度のケースについて示したものである。予備校における偏差値が合格ライン近辺の学生では、一発試験においては失敗するケースがあり、不合格となってしまっているケースも多い。ただし、従来のように入試が高倍

率であった時代には、この重なりの問題はあまり問題視されない。倍率が下がってくると、両分布は先の図に示したように近づきあい、合格者の平均偏差値に近い学生でもかなりの数が不合格となってくる。このような状況下では、如何に真の学力が高い学生を正しく検出して合格にさせるような試験方法の構築が重要となる。例として、合格者・不合格者の平均偏差値が同じで、分散が異なる場合を以下に比較してみる。図2と図3に比較するように、合格者分布と不合格者分布の双方で標準偏差が比例して変化した場合、両社において難易度レベルに変化が生じない。図3のような選抜が行われてしまうと、体外的な難易度には影響がなくとも、入学後の学生のレベルに大きなばらつきがみられるようになり、大学教育自体に支障をきたす可能性がある。大学入試においては、単純な偏差値による難易度だけでなく、合格者と不合格者の偏差値分布で評価することが重要と考えられる。

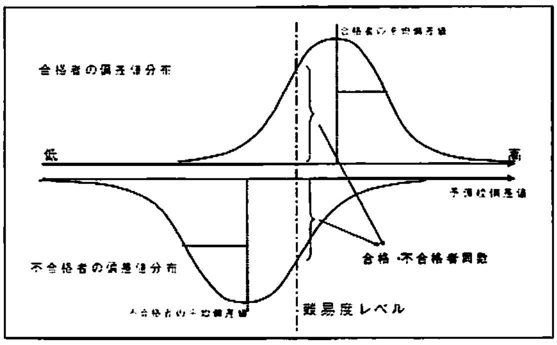

図2．偏差値分布と難易度レベル（分散が大きい場合）

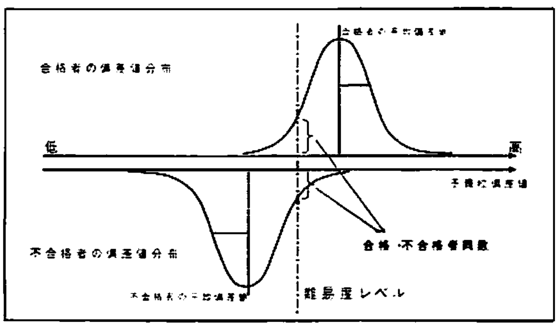

図3．偏差値分布と難易度レベル（分散が小さい場合）

次に、入学試験の精度を高め、一発試験の偶然性を極力排除して、本来学力の高い学生を正しく選抜することに努めるというケースを考えてみる。この場合、先の分布概念図において、全体の受験者数を固定とし、不合格者のうちの上位を合格者の下位と入れ替える操作とみなすことができる。この操作は、確実に難易度レベルを上昇させる。

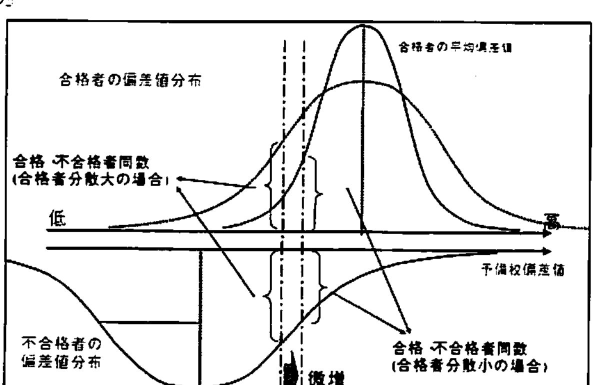

図4. 合格者の標準偏差を小さくした場合の挙動

すなわち、様々な施策によって優秀な受験者を集める努力は重要であるが、入試の精度を上げることによっても偏差値（難易度レベル）は向上する。

ただし、本来、学生各々の真の学力は実際には測定することができず、これも予備校などの模擬試験を受験した際の偏差値に基づいて推定しなければならない。すなわち、受験者の偏差値分布のばらつきは"受験者の学力分布"に"模擬試験誤差"が加算されたものであることを考慮しておく必要がある。

4. 複数回入試に関する検討

近年の大学市場競争の激化を受け、同一学科の入学試験を複数日に渡って実施するケースが見られるようになった。一般にこのような試験方法は、偏差値で言うところの中堅以下の大学で多く見られ、しばしば受験生確保の方策と見なされることから市場評価を下げるとの批判もある。しかし一方で、複数試験による評価は試験結果の信頼性を向上させることも期待できる。とくに、倍率が低下してきた場合には、先の章で示したように合格者と非合格者の学力レベルが拮抗し、単発の試験による運不運による入学者選別結果への影響は大きくなっている可能性がある。本章では、統計的に検出力を高める方策としての複数回の入学試験を実施する方策について検討を行う。

ここでは、マクロな視点に立ち、受験生は真の学力に従って、合格とすべき群と不合格にすべき群に分けられ、合格群と不合格群はそれぞれ正規分布に従うモデルを考える。不合格者の合格者に対する比をδとおけば、この入学試験の実質倍率は$\delta+1$で与えられる。

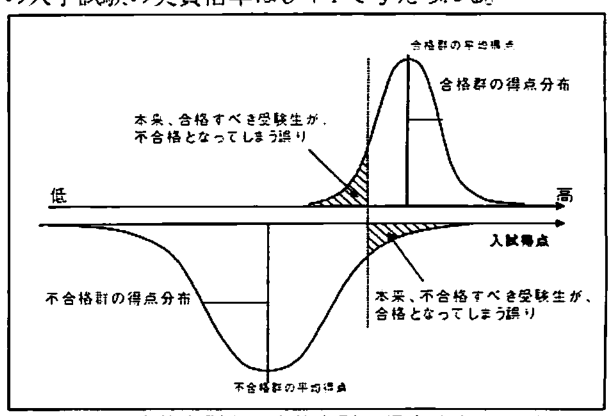

図5．合格者群と不合格者群の得点分布の関係

ここで注意すべき点は、合格群とは"本来、合格すべき受験生のグループ"を表しており、実際に試験で合格する受験生群とは異なる点である。合格群の確率分布を$N(\mu_1, \sigma_1^2)$、不合格群の確率分布を$N(\mu_2, \sigma_2^2)$とする。このとき、1 回のみの試験によって得られる受験者得点分布は、図5のように与えられる。不合格者の対合格者比δであることから、不合格者の得点分布の面積は、合格者の得点分布の面積のδ倍になっている。また、合格者の人数は一定であると考えれば、本来合格すべき受験生が不合格となってしまう誤り部分の面積と、逆に本来不合格にすべき受験生が合格となってしまう誤り部分の面積は等しくなる。

ここでは、複数の試験日を設けるモデルについて、最も基本的で単純なケースで本質を考察するため、複数回に渡って同一の受験者が前日受験し、かつそれらの得点の平均点を用いて入試判定を行うという特殊なケースを考える。同じ受験者が複数の受験日に渡って受験するケースは、とくに志望の強い大学学部においては一般的に見られるケースである。このとき、一般にn回の試験を実施し、その平均値をもって入試判定を行うという単純なモ

デルで記述することができる。合格判定ラインを入試得点で K 点とおく。このとき、先に述べた2つの誤り人数の同値性から、

$$\Pr\left\{Z > \frac{K-\mu_1}{\sigma_1/\sqrt{n}}\right\} = \delta \cdot \Pr\left\{Z < \frac{K-\mu_2}{\sigma_2/\sqrt{n}}\right\}$$

という関係式が導かれる。ここで、$\Pr\{Z < u\}$ は標準正規分布に従う確率変数 Z に対し、$Z < u$ となる確率を表す。とくに、実質倍率が2倍のケース、すなわち、$\delta = 1$ の場合については、K について明示的に解くことができて、

$$\frac{K-\mu_1}{\sigma_1/\sqrt{n}} = \frac{\mu_2 - K}{\sigma_2/\sqrt{n}}$$

という関係式から、

$$K = \frac{\sigma_1\mu_2 + \sigma_2\mu_1}{\sigma_1 + \sigma_2}$$

という結果を得る。すなわち、試験回数 n に対して合否判定のラインは不変である。このとき、

$d = \dfrac{\mu_1 - \mu_2}{\sigma_1 + \sigma_2}$ という合格者平均得点と不合格者平均得点の、標準偏差に対する基準化得点差が重要な意味を持つ。本来、合格者とすべき受験生を不合格としてしまう誤りは、逆の不合格者とすべき受験生を合格としてしまう誤りと等しく、$\Pr\{Z > \sqrt{n}d\}$ で与えられる。ここで具体例として、$\mu_1 - \mu_2 = 100$ とし、いくつかの $\sigma_1 + \sigma_2$ に対して、試験回数 n を変化させたときの誤りについて考えてみる。表1に、実際に正規分布の上側確率から求めた誤り確率の値を示す。

この結果、合格者と不合格者の各分布の標準偏差が、両者の平均得点差 $\mu_1 - \mu_2$ と同程度の標準偏差を持つとき、標準偏差和である $\sigma_1 + \sigma_2$ の値は2倍程度の値を持つことになり、1回の試験で合格者を決定すると、30.9%程度の誤り率を持つことが分かる。合格者と不合格者の各分布の標準偏差が、両者の平均得点差 $\mu_1 - \mu_2$ の2倍程度の値を持つときには、標準偏差和である $\sigma_1 + \sigma_2$ の値は4倍程度の値を持つことになり、このとき誤り率は40.1%にもなる。すなわち、4割もの合格にすべき優秀な受験者を逃してしまうことにつながる。また、試験回数 n を増やし、入試の手間を増大させることで、試験の誤り率を低減させる施策をみても、$\sigma_1 + \sigma_2$ の値が大きいほど、期待ほどの効果は得られない。このことを分かり易くするために、各々の平均得点差 $\mu_1 - \mu_2$ と標準偏差和 $\sigma_1 + \sigma_2$ に対し、試験回数 n を増やした時の誤り率の試験回数 n が1である通常の1発試験に対する比を示したのが表2である。$\mu_1 - \mu_2$ と $\sigma_1 + \sigma_2$ が同じ場合には、2回の試験を実施することによって、誤り率は約半分の49.6%まで低減させることができるが、試験結果のばらつき指標である標準偏差和 $\sigma_1 + \sigma_2$ が $\mu_1 - \mu_2$ の2倍であるときには、77.7%までしか低減させることができない。すなわち、1回の試験の精度を高めておかないと、試験日を複数にすることの意味が薄れてしまうのである。

表1．平均差と標準偏差和と試験回数による誤り確率の関係（$\delta = 1$ の場合）

平均差 $\mu_1 - \mu_2$	標準偏差和 $\sigma_1 + \sigma_2$	試験回数 n 1	2	3	4	5
100	20	2.87×10^{-7}	7.74×10^{-13}	－	－	－
100	50	0.023	0.002	2.66×10^{-3}	3.17×10^{-5}	3.88×10^{-6}
100	80	0.106	0.039	0.015	0.006	0.003
100	90	0.133	0.058	0.027	0.013	0.006
100	100	0.159	0.079	0.042	0.023	0.013
100	200	0.309	0.240	0.193	0.159	0.132
100	300	0.369	0.319	0.282	0.252	0.228
100	500	0.421	0.389	0.365	0.345	0.327

表2．試験回数の増加による誤り確率低減度の割合（$\delta = 1$ の場合）

平均差 $\mu_1 - \mu_2$	標準偏差和 $\sigma_1 + \sigma_2$	試験回数 n 1	2	3	4	5
100	20	1.00	2.70×10^{-6}	－	－	－
100	50	1.00	0.103	0.012	0.001	1.70×10^{-4}
100	80	1.00	0.365	0.144	0.059	0.025
100	90	1.00	0.436	0.204	0.099	0.049
100	100	1.00	0.496	0.262	0.143	0.080
100	200	1.00	0.777	0.626	0.514	0.427
100	300	1.00	0.863	0.763	0.683	0.617
100	500	1.00	0.924	0.866	0.819	0.778

5. まとめ

本稿では、大学入試の基礎的な側面について、確率分布モデルの考えから議論を行った。このような議論は、しばしば入試方式の変更が受験生に与える印象やPR効果などの側面を捉えておらず、本質的な数理的インパクトのみを考えているものであるが、入試戦略の立案において念頭に入れておくべき知見を与えていると考えられる。とくに、入学試験の判定精度の向上は、その他の入試施策の有効性に影響を与えることが示唆され、単に入試戦略を検討するだけでなく、一つ一つの試験について優秀な学生を正しく判定できるように精度を高めることの重要性が明らかとなった。

参考文献

[1] 大学入試センター，http://www.dnc.ac.jp/index.htm

[2] 佐藤忠彦："大学におけるマーケティング"，筑波フォーラム，No.71, pp.113-116, (2005)

[3] 柳井 晴夫，前川 眞一：大学入試データの解析，理論と応用，現代数学社，(1999)