大学入試得点と模擬試験得点の統計モデルおよび入学手続率の分析†

高橋静昭†† 安原治機††† 長嶋祐二††

模擬試験の得点や共通一次の自己採点などから，入試における合格最低点を推定する場合，合格者の学力の水準と入学の難しさの2つは，合格最低点とは微妙に異なる．それは，模擬試験の得点等と大学入試得点の相関係数の値が1より小さいことに起因する．もし相関係数が1ならば，合格最低点の入試得点者のみが，最低点に対応する模試の境界得点者となる．しかし，相関係数が1より小さくなるに伴って，合格最低点よりやや高い入試得点者と低い得点者の双方が模擬試験の対応する境界得点となることが多くなる．合格率が1/2以下のときには，高得点者の方が少ないので前者の人数が後者を上まわる．したがって模擬試験の対応境界得点近傍の受験者の合格率は，1/2より小さくなる．さらに同様の関係で，合格最低点近傍の受験者の模試の平均点は模試の境界得点より低くなる．逆に，合格率が1/2を超えているときは，模試の境界得点の受験者の合格率は1/2より大きくなる．これらの関係は，従来公表されている難易度では十分に説明されていない．本論文では，模試得点や自己採点等の変数 X とそれによって評価される大学入試得点 Y が，2次元のガウス分布に従うものとしてモデル化し，合格最低点，最低合格者の水準および合格難易度の3つを定量的に定義し検討する．さらにこの統計モデルを用いて，合格者の中の入学手続者の入試得点上の分布を表現し，入学手続の予測に応用することを提案する．

1. はじめに

各大学の入試制度と高等学校の大学受験指導，および両者の間にあって入試情報を分析し提供する受験産業さらに文部省等の行政指導も加わって，我が国の大学入試制度はシステム化されている．この現在のシステムに対して，偏差値（成績を平均と分散に対して基準化した値）の過度な偏重その他の様々な批判がある．しかし，各大学がこのシステムの力学を無視して独自に行動することは困難な面がある．例えば工学部の勉学には物理と化学の両方が必要であるとの判断に基づいて，ある工科系単科の私立大学で，両方の基礎を含む「理科1」を必修にしたところ，受験勉強の負担増を嫌って，受験者が大幅に減少した．全国的な入試システムに組み込まれているという立場は，高等学校も受験産業も同様である．

他方そのような枠組みの中で，大学は入試を通じて，自由市場の原理に基づく企業間と同様の競争を演じている．したがって，企業秘密と同様に公開できない情報もあり得る．しかし，たとえ私学といえども公的責任もあり，受験者数や選抜方式その他の事柄を公開する義務がある．また，他機関と情報を交換し合い，円滑な大学運営に役立てる必要もある．

合格難易度（大学の学部や学科ごとに入試で合格する難しさの程度を示す値）の研究は，現在の偏差値偏重の行き過ぎをさらに助長するのではないかとの心配は無用である．試験による学力の評価には，ある一定の誤差を伴うことが経験的に確かめられる．その誤差の標準偏差は，受験者の学力分布の標準偏差の半分程度である．それより細かい値を用いた大学の偏差値の算出は，計算の上では可能ではあっても意味がない．本研究は偏差値偏重の行き過ぎを批判する材料になる．

さらに，合格難易度を検討するためにここで用いる統計モデルが，合格者の入試得点上の入学手続者の分布のモデル化にも役立つことが示される．各受験者が複数の大学を受験して，合格する大学が1人平均3校あるとき，大学側からみれば合格者の3人に1人しか入学手続をしないことになる．入学手続者数の合格者数に対する比率を手続率という．国立大学ではこの手続率がほぼ100% に近いが，私立大学では 30～40% 程度となる大学も多い．また，手続率は大学の学科や年度ごとに特性が異なり，正確な予測が難しく，私立大学にとっては経営の根幹に係わる重大関心事である．

2. 大学ランキングの従来の評価法

受験産業が企画し実施している大学入試模擬試験

---

† Relationship between Marks of Entrance Examination to University and of Sham One by SHIZUAKI TAKAHASHI (Department of Electronic Engineering, Kogakuin University), HARUKI YASUHARA (Department of Architecture, Kogakuin University) and YUJI NAGASHIMA (Department of Electronic Engineering, Kogakuin University).

†† 工学院大学電子工学科
††† 工学院大学建築学科

や，あるいは国立共通一次の自己採点（試験後，各受験者自ら自分の解答を思い出して採点したもの）を用いた，各大学の合格難易度等が公表されている．これらは，受験産業のアンケート調査から得られる各受験者の合否の情報に基づき，模試得点または共通一次自己採点の得点上の合格者の分布から計算される．しかし手続率を考えると，大学ごとの合格者の平均得点は必ずしも合格者の水準を示しているわけではない．

大学の合格難易度として，次の評価方法[1]がある．N_1 人の合格者と N_2 人の不合格者があり，それぞれの模試または自己採点の平均点を P_1，P_2 とするとき，合否の人数の逆数による重み平均 A を用いる方法である．すなわち，

$$A=\{(1/N_1)P_1+(1/N_2)P_2\}/\{(1/N_1)+(1/N_2)\}$$
$$=n_2P_1+n_1P_2 \qquad (1)$$

ただし，n_1，n_2 は，それぞれ合格率と不合格率であり，$n_1=N_1/(N_1+N_2)$，$n_2=N_2/(N_1+N_2)$ とする．

この値 A は次のように考えると，入学試験における合否の境界点の推定値とみなすことができる．大学入試の得点を，模試得点等と平均値が等しくなるように換算して考える．さらに粗い近似ではあるが，入試得点上の受験者数密度を一様と仮定し，入試得点が境界値 A 以上の合格者の平均を P_1'，A 以下の不合格者の平均を P_2' とすると，**図1**のように示される．さらに P_1'，P_2' の推定値として，合格者と不合格者の自己採点のそれぞれの平均 P_1，P_2 を用いると，A の値として式(1)が導かれる．

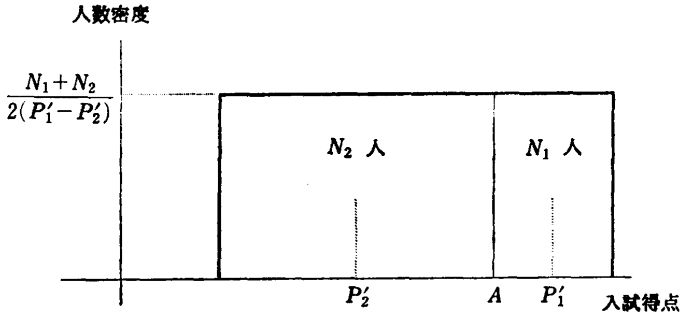
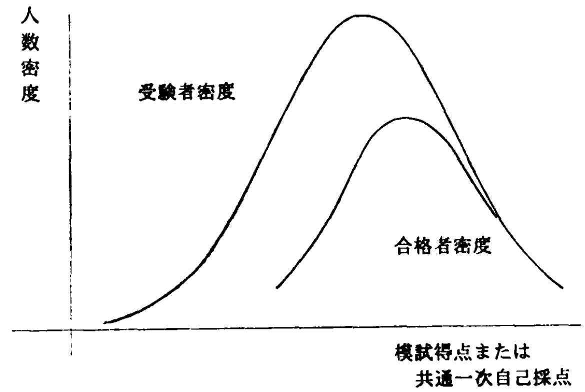
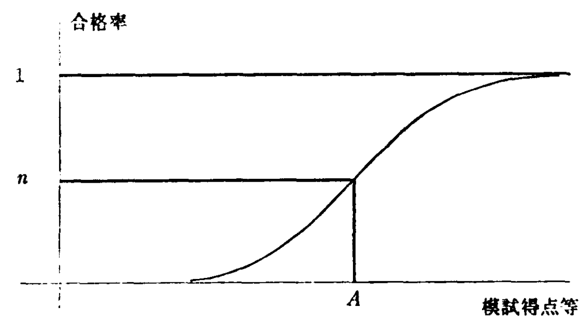


図 1 単純化した合否分離モデル
Fig. 1 Simplified separating model.

もう一つの合格難易度の評価方法として，模試得点上の合格率曲線から定める方法[2]もある．一般に，受験者の模試得点分布は**図2**(a)に示すように，ほぼガウス分布になる．そこで，得点上の局所的合格率が，その大学の全体の合格率に等しい模試得点を合格難易度と定める．この値は，確かに入学の難しさの程度を評価している．しかし，大学ごとに異なる合格率を用いて評価することに疑問が残る．



(a) 受験者および合格者分布
Distribution of applicants and successfuls.



(b) 合格率分布
Characteristic of passing rate.

図 2 得点上の合格者分布の特性
Fig. 2 Relationship between mark and passing rate.

ここで，図2(b)に示す合格率曲線を実際にプロットする際には，得点を等間隔にクラス分けしてはならない．模試得点分布の両裾では人数密度が少ないので，たまたまそのクラス内に入った 1，2 名が全員合格や不合格になることがあり，滑らかな局所適合確率の曲線を描けない．中央と両裾における合格率の値の分散を同程度にするため，合格者数を等しくするように両裾では広い得点間隔にする必要がある．

3. 合格最低点の推定

模擬試験の得点 X と大学入試得点 Y を個人ごとにつき合わせたデータを持っている機関はない．ここでは模試得点 X とその個人の大学入試の合否のデータから，入試得点 Y との関係を推定することを考える．模試得点 X と，入試得点 Y は2次元のガウス分布に従うと考えられる．そこで，定数 α，β によって $y=$

$\alpha+\beta Y$ と1次変換をして，y の平均と分散を X のそれらと等しくすれば，X と y は対称な2次元のガウス分布に従う．したがって**図3**のように，確率密度の等高線は長軸が45度傾いた楕円となる．この α,β は大学ごとに定められるべき定数であるが，値は未知でありかつ直接知ることはできない．大学が $y\geqq A$ の受験生を合格にしたとき，図2(a)のような合格者密度分布となる．確率密度が X と y に対して対称なので，$y\geqq A$ の人数と $X\geqq A$ の人数はほぼ等しい．したがって，模試得点の上位から合格者数だけ数えれば，最後の人の得点が大学入試得点を模試得点に換算したときの合格最低点の推定値となる．

4. 合格難易度と合格者レベル

計算の簡略化のために，平均値0，分散1に正規化した X と Y の値を u,v とする．ρ を X と Y の相関係数とすれば，仮定により，ρ は u と v の共分散でもある．また，u,v の同時確率密度 f は次式で示される．

$$f(u,v)=\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{u^2-2\rho uv+v^2}{2(1-\rho^2)}} \qquad (2)$$

さらに，$E(X)=P$, $Var(X)=S^2$ とすると，合格最低点 A に対応する v の値 a は

$$a=(A-P)/S \qquad (3)$$

である．全受験者数を N，合格者数を N_1，合格者の模試の平均点を P_1 とし，対応する u の値を p_1 とする．正規化模試得点と入試点の値が，それぞれ u,v の近傍における人数密度は $N\cdot f(u,v)$ である．模試得点 u と人数密度の積を $u\,v$ 平面上の合格者の領域 $(a\leqq v)$ で積分して，全合格者の模試得点 u の合計の期待値 $N_1\cdot p_1$ を示すと次式となる．

$$N_1\cdot p_1=\int_{-\infty}^{\infty}\left\{\int_a^{\infty}uNf(u,v)dv\right\}du \qquad (4)$$

式(4)の右辺の積分の順序を逆に計算して，次式が得られる．

$$\begin{aligned}N_1\cdot p_1 &=\int_a^{\infty}\frac{N}{\sqrt{2\pi}}e^{-\frac{v^2}{2}}\left\{\int_{-\infty}^{\infty}\frac{u}{\sqrt{2\pi(1-\rho^2)}}e^{-\frac{(u-\rho v)^2}{2(1-\rho^2)}}du\right\}dv\\ &=\int_a^{\infty}\frac{N}{\sqrt{2\pi}}\rho ve^{-\frac{v^2}{2}}dv=\frac{1}{\sqrt{2\pi}}N\rho ve^{-\frac{a^2}{2}}\end{aligned}$$

これから，ρ が次式で定まる．

$$\begin{aligned}\rho&=\sqrt{2\pi}n_1p_1e^{\frac{a^2}{2}}\\ &=\sqrt{2\pi}N_1(P_1-P)e^{\frac{(A-P)^2}{2S^2}}/(NS)\end{aligned} \qquad (5)$$

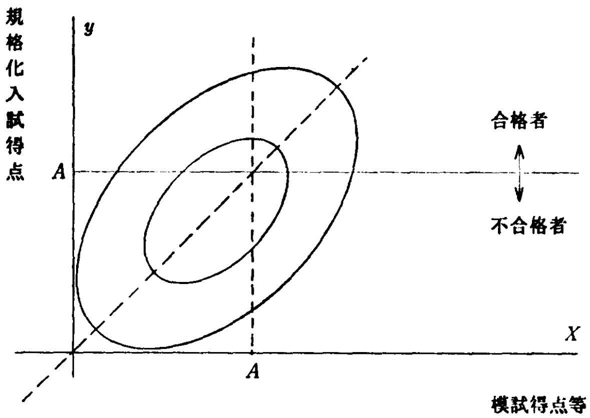


図3 Xy 平面
Fig. 3 Xy plane.

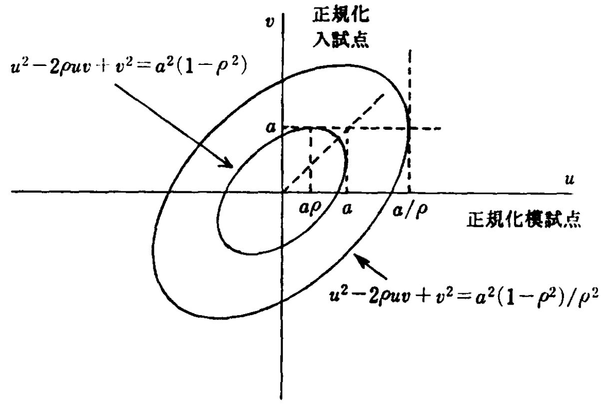


図4 正規化模試点と正規化入試点
Fig. 4 Normalized marks of sham examination and actual one.

式(5)で用いられている変数の中で，大学ごとの模試得点と合否のデータのみから直接には得られない値は ρ と A のみである．したがって，3章の A の推定値を代入することにより，ρ もまた推定できることになる．

uv 平面上の等確率曲線は楕円の曲線群であり，定数 k を用いて，$u^2-2\rho uv+v^2=k^2$ で示される．各 u の値における v の条件付確率密度は，ガウス分布となる．そして，**図4**に示すように，u を一定とする直線と楕円の接点の v 座標が確率密度の最も高い得点で，v の条件付確率密度分布 $p(v|u)$ の中心となる．楕円 $u^2-2\rho uv+v^2=a^2(1-\rho^2)/\rho^2$ と直線 $u=a/\rho$ は，図4に示すように $v=a$ で接する．したがって，$u=a/\rho$ の人が $v\geqq a$ となって合格する条件付確率は1/2である．合格の可能性として1/2の確率が与えられるのに必要な模試得点 a/ρ をここでは合格難易度と定め，以下で

はこの定義に従う．

逆に，$v=a$ の人の u の条件付確率 $p(u|v=a)$ は $u=a\rho$ を平均とするガウス分布となることが，図4から分かる．ここでは，$v=a$ に相当する合格最低得点者の模試得点 u の平均点 $a\rho$ を合格者レベルと定める．

これまでに定めた合格難易度，合格最低点，合格者レベルの評価式 a/ρ, a, $a\rho$ の3つは，それぞれの利用目的によって選択されるべきであり，異なる意味を持っている．大学のレベルを単純に比較する場合には，合格最低点 a が適当であるが，模試得点のみから合否の可能性の程度を知りたいと思う受験生にとっては合格難易度 a/ρ は重視すべきデータである．模試で評価される要素以外の実力を大学が評価する際に，相関係数 ρ は小さくなる．このとき，模試得点の低い人が他の実力を評価されて合格されやすくなり，模試得点が a 点の人の合格率は減少し，合格率 1/2 を得るための模試得点すなわち合格難易度は高くなる．その分，合格最低点の人の中には，模試得点の低い人が他の実力が評価されて合格するので，合格最低点の人の模試得点による評価すなわち合格者レベル $a\rho$ は低くなる．しかし，この $a\rho$ を大学のレベルと評価するのは適当ではないように思われる．

相関係数 ρ は大学受験体制のメカニズムの中で，均衡を保って，ある値に落ち着くようになっていると思える．ρ の値があまり1に近いと，模試低得点者の合格の可能性が少なくなるので，受験しなくなる．受験者の模試得点の分散が小さくなると，ρ は減少する．逆に，ρ が極端に小さいとその年の受験者層から外れた得点者も合格の確率があるので受験するようになり，ρ は次年度には増加する．

以上の検討を含めて，相関係数 ρ が小さくなる理由として，以下の事柄が列記される．

(1) 筆記試験による学力調査が本質的に有する誤差
(2) 模試から入試までの期間の勉学効果
(3) 受験生の体調の影響
(4) 入試問題，配点，採点基準における大学の個性
(5) 内容の偏り等による問題の不適切
(6) 受験生の模試得点分布の狭さ

理由(4)による ρ の減少は，意識されていれば問題ではない．

項目(1)の誤差の程度に見当を付けるため，模試と入試に同程度の誤差を含むとした場合を検討する．個人の真の学力を g とし，誤差 e_X と e_Y が加わって，$X=g+e_X$, $y=g+e_Y$ となっていると考える．g, e_X, e_Y は互いに独立である．X と y の分散を S^2 とし，e_X と e_Y の分散を σ^2 とすれば，$Var(g)=S^2-\sigma^2$ となる．また，$\rho=Cov\{X,y\}/S^2=Var(g)/S^2$ であるから，

$$\sigma/S=\sqrt{1-\rho} \tag{6}$$

となる．相関係数 ρ の値が 0.7〜0.8 程度のとき，相対誤差 σ/S は 0.45〜0.55 程度となる．この場合，入試における標準偏差の半分程度が，試験固有の標準偏差となる．

5. 公表されている数値に基づく計算例

昭和57年度の北海道大学（理Ⅰ系）の共通一次の自己採点にみる合否の状況に関して，旺文社が行ったアンケート調査の結果を表1[1]に示す．自己採点の分散 S^2 が示されていないので，各階級の中央の値の人がその度数だけいると考えて，$S^2=75.2^2$ を計算した．また，受験者平均 P が，示されている値725になるように，500点以下については418を代表値とした．自己採点に換算した合格最低点 A は，$A=729$ と推定される．その他 $N=619$, $N_1=324$, $P_1=771$ 等を式(5)に代入すると，自己採点と入試得点の相関係数 ρ が求まり，$\rho=0.8035$ となる．合格難易度すなわち合否が半々以上となるために必要な自己採点を A_d とする．また合格者レベルすなわち合格最低点近傍の人の自己採点の平均を A_l とする．これらを規格化した値がそれぞれ a/ρ と $a\rho$ である．数値を求めると，$A_d=730$, $A_l=728$ となる．$A>P$ のときは図4に示すように，$A_l<A<A_d$ となるが，$A<P$ の場合には $A_l>A>A_d$ となる．この例では，A と P がほぼ等しく，これらの数値が，すべて近い値を取っている．

6. 入学手続者密度分布

この章では大学の立場から，合格者の中の入学手続者の入試得点上の分布を検討する．大学が持っているデータは入試得点とその個人が入学手続をしたか否かの情報である．これまでは，規格化模試得点 u から規格化入試得点 v を吟味した．この章では，直接知ることのできない辞退志向度を想定し，それを入試得点で吟味する．入試得点の記号がこれまでとは逆になるが，入試得点と辞退志向度のそれぞれの規格化変数を改めて，u, v とすれば，これまでの検討結果を，その

まま新しい問題に適用できる．

入試点 u と辞退志向度 v が，相関係数 ρ の2次元ガウス分布に従っているものとする．さらに図5に示すように，ある定数 b に対して $v \geqq b$ の人は合格しても入学を辞退するものとする．b が大きいほどその大学の人気は高いものと考えられる．正規化入試得点 u における局所的手続率は次式で示される．

$$l(u)=\int_{-\infty}^{b} Nf(u,v)dv \Big/ \int_{-\infty}^{\infty} Nf(u,v)dv$$
$$=F\{(b-\rho u)/\sqrt{1-\rho^2}\}$$
$$=F(G-Ru) \qquad (7)$$

ただし，$F(x)$ はガウス分布の分布関数で，さらに

$$R=\rho/\sqrt{1-\rho^2}$$
$$G=b/\sqrt{1-\rho^2} \qquad (8)$$

とする．この G を手続指数と呼ぶことにする．G は大きいほど手続率がよい．

私立大学における通常の R の値は経験的にほぼ1に近い．すなわち多くの場合，手続率曲線 $l(u)$ は入試得点上位者から u までの受験者累積曲線を平行移動したものとなる．このとき，$\rho=1/\sqrt{2}\fallingdotseq 0.7$ である．$l(u)$ は近似的に次式となる．

$$l(u)=F(G-u) \qquad (9)$$

表1　北大（理Ⅰ系）の共通一次にみる合否状況
Table 1　The state of common preliminary of entrance examination to Hokkaido University.

自己採点合計	合格者数	不合格者数
1000～900	0	0
890	1	0
880	3	0
870	1	0
860	2	0
850	4	1
840	9	0
830	6	0
820	18	0
810	21	1
800	21	0
790	25	0
780	30	2
770	22	4
760	30	10
750	27	12
740	23	11
730	26	12
720	11	21
710	14	24
700	11	25
690	9	19
680	5	17
670	2	20
660	1	14
650	2	16
640	0	10
630	0	10
620	0	10
610	0	11
600	0	8
590	0	6
580	0	4
570	0	4
560	0	7
550	0	3
540	0	2
530	0	1
520	0	2
510	0	2
500～0	0	6
総人数	324	295

合格者平均：771点，不合格者平均：675点，アンケート回答者以外を含めた全合格者数：324名，同じく全不合格者数：554名

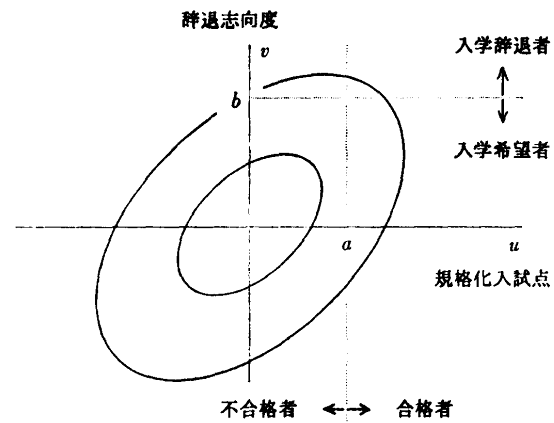


図5　入学辞退者分布の統計的表現
Fig. 5　Statistical model of distribution of refuser.

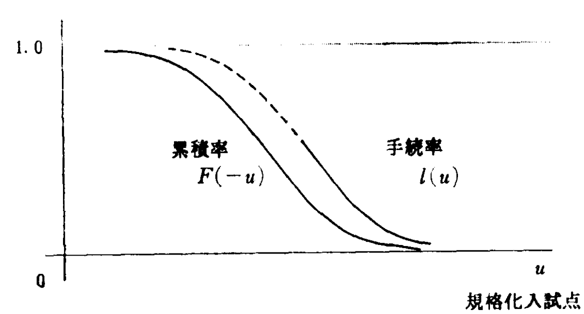


図6　手続率と受験者累積率曲線
Fig. 6　Entrance ratio and accumulation of applicants.

したがって，入試得点上に上位者からの受験者累積率曲線 $F(-u)$ と手続率曲線 $l(u)$ を図示すると図6のようになる．このモデル化に従えば，$\lim_{u\to\infty} l(u)=0$ となるが，得点 u の高い人でも，他に受験したすべての大学が不合格になる確率は0ではない．これは，受験したすべての大学が不合格になっても入学して来ない受験生のいることを意味する．これは矛盾する結果のように感じられるかもしれないが，事実である．$u \geqq a$ のすべてを合格とするとき，全合格者数は $N\{1-F(a)\}=N\cdot F(-a)$ となり，$u \geqq a$ の全体の手続率 $r(a)$ は N に無関係となり，

$$r(a)=\int_a^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}F(G-u)du/F(-a) \tag{10}$$

で示される．したがって，手続率 $r(a)$ は合格率 $F(-a)$ と G によって，一義的に定まる．この関係を図7に示す．この F-r 曲線を用いると，受験者数の異なる入試状況においても手続の程度を他の要因から分離して比較することができる．しかし，累積合格者数と累積手続者数のグラフの方がデータの分散が小さく，実際にデータを分析する際には，グラフを描きやすい．

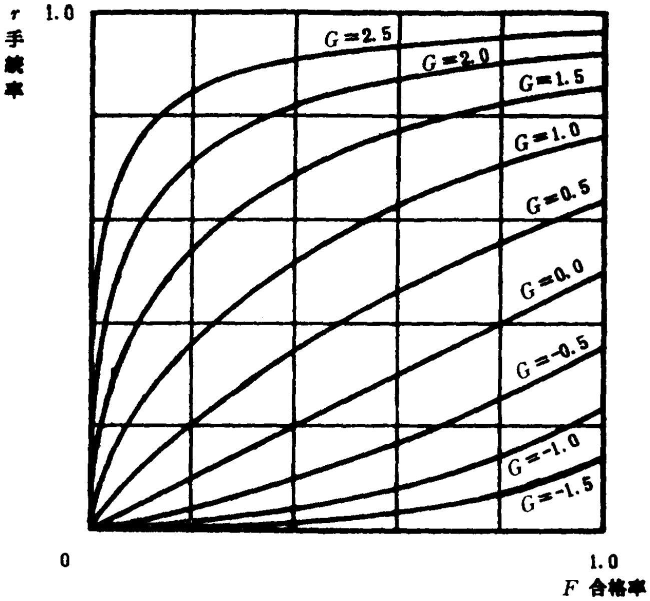


図7 $R=1$ とするときの合格率と手続率の関係
Fig. 7 Relationship between passing ratio and entrance one in the case of $R=1$.

入試得点と辞退志向度の相関係数 ρ は，模試と入試の場合と同様に次の理由で減少し得る．ρ が小さいときには手続率曲線の勾配は緩やかになる．

（1） 学力調査に本質的に含まれる誤差
（2） 他大学の合否の基準の変動
（3） 受験生の体調等の影響
（4） 問題，配点および採点基準における大学の個性
（5） 内容の偏り等による問題の不適切性

7. むすび

ここでは，合格最低点と合格難易度および合格者レベルの三者の違いと，そのメカニズムを明らかにした．また，合格者の中の入学手続者分布を示す統計的モデルも提案した．これらの推定は，みな，点推定である．区間推定については今後の課題である．また，式(8)の R の値が工学院大学では多くの場合ほぼ1となるが，それには理論的根拠があり他の私立大学でも同様なのか，あるいはそうではなく異なるのかどうかについても，今後検討されなくてはならない．

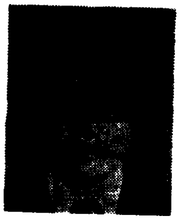
謝辞 本研究について，多大の御協力を頂いた入試広報課の方々，および常に励ましてくださった工学院大学電子工学科の長嶋秀世教授に深謝する．

参考文献

1) 全国大学合格難易ランキング，螢雪時代，昭和57年7月号第1付録，旺文社，東京 (1982).
2) 全国大学合格難易ランキング，「KEISETSU 螢雪時代」，昭和58年7月号特別付録，旺文社，東京 (1983).

（昭和62年1月8日受付）
（昭和62年9月9日採録）

髙橋 静昭（正会員）
昭和16年生．昭和39年工学院大学工学部電気工学科電子工学コース卒業．昭和41年同大学院修士課程修了．同年同大学助手．昭和46年同大学講師．昭和54年同大学助教授．工学博士．現在，学習回路モデルの記憶効率および入試統計分析に興味をもっている．電子情報通信学会，画像電子学会各会員．



安原　治機

昭和18年生．昭和42年工学院大学工学部建築学科卒業．昭和44年同大学院修士課程修了．同年同大学助手．昭和61年同大学講師．現在，住宅デザインに関する統計，情報科学を通しての建築設計方法（建築 CAD 等）および大学計画研究に興味をもっている．著書「調査方法と分析方法」（共著，日本建築学会），「パソコン CAD・CG 用タブレット／デジタイザーの使い方」（槇書店）「建築・都市計画のための調査・分析方法」（共著，井上書店）．日本建築学会会員．

長嶋　祐二（正会員）

昭和28年生．昭和53年工学院大学工学部電子工学科卒業．昭和55年同大学院修士課程修了．同年同大学助手．現在，光強度分布直視装置システム，数値計算法および画像処理の研究に興味をもっている．著書「マイコンによる数値計算法」（共著，昭晃堂）．ACM，電子情報通信学会，テレビジョン学会各会員．