入学試験の統計精度について

服部　哲弥
宮沢　弘成

１．まえがき

試験による選抜は社会生活上重要な場面で頻繁に行われる．社会的に重要な一つの例として大学の入学試験がある．宮沢（行動計量学15巻1号(1987)34-42）は大学入試に関して，試験成績が正しく受験生の能力を測っているかどうかという点に注目し，統計学的な側面を一つの中心として研究を行った．その際，入試の統計精度を特徴づける量として入学試験の正確度という概念を定義した．入試の正確度は入試の「誤差」の小ささを反映する量なので，入学試験の適切な実施方式を決定するための合理的基準となりうるものである．

この論文の主要な目的は正確度に関してより詳細な計算・客観的分析を加えることにある．その帰結として正確度が入学試験の方式決定に役立つことを示す典型的な結果も併せて紹介したい．ここでの分析によって，例えば以下のような問いに答えることができる：

問題１：２段階選抜（いわゆる足切り）方式の入学試験において全倍率が決まったとき，１（２）次試験の倍率はどう選ぶのが最善か？

問題２：２回分の試験問題がある．これを足切り方式で利用するのと，分離分割方式で利用するのはどちらがよいか？

問題３：全倍率が決まっているとする．N 回の多段階選抜試験において，１回毎の試験が誤差を含んでいても N を大きくすれば全体の誤差は0に近づくか？

２．入試の誤差と正確度

入学試験を受験生の能力の測定ととらえる．即ち，入学試験を行う目的は大学での勉学のために望ましいと考えられる能力をより多く備えている受験生から合格させることである，と考える．この能力を測定すべく然るべき試験問題をいくつか用意したとき，その成績得点は一定ではなく試験問題毎に変動する．変動の原因は，偶然その問題の類題をやっていたかどうか，試験時の体調，そのほか種々考えられるがその子細はこの論文の分析には影響しない．重要なのは，個々の試験問題の成績が変動する故に受験生の能力を測定するためには無限の時間（無限個の問題）を用意しなければならないという点である．入学試験に即して言うならば，能力を多く備えている受験生から順に定員まで合格させるのが理想であるにも関わらず，それを完全に実行するには無限に長い入学試験を行わなければならないということである．

もちろん有能かつ将来性ある受験生の能力と時間を入試という測定だけに費やすのは貴重な人的資源の浪費である．入学試験はできるだけ短く（受験生の労力を最小に）するのが望ましい．即ち現実の試験は有限の時間で測定を行う．従って測定誤差が伴う．言い替えると，入試をやり直すと合格者には殆ど必然的に入れ替わりが生じる．そこである入学試験を行うとき，その試験の誤差がどれくらいあるか，誤差を小さくするためにはどのような方式を取るべきか，が重要な問題となる．この論文で取り扱うのはこの問題である．

注意１：入試は個人の能力を測定するが，大学においてどの能力を必要とするかは，学校・学部・社会状況など種々の要因によって変わり得る．例えば広い意味での計算能力を必要な能力とする考え方，別の例として広い意味での知識を能力

とする考え方，がある．欲しい人材と試験問題が整合していればよく，以下ではどのような能力を要求するかとは独立に議論できる部分を論じる．
注意２：人材の能力を指定したとき，その能力を測定する試験問題が作れるかどうかは難しい問題である．例えば，仏教伝来の年代を問うのは知識を測るが，同時に記憶能力の測定にもなるであろう．他方数学を公式でなく考えて解かせるのは大脳の情報処理脳力性能測定に近いだろう．この論文で計算した諸結果を実際に適用するためには適切な入学試験問題が必要である．そのような「然るべき」試験問題を実際に作り以下の計算結果を検証していくことは難しい問題ではあるが可能であると期待しよう．これは入学試験を科学的に研究する際の実験の役割を果たすであろう．

以下，入試における統計精度（測定誤差）の問題に集中して，基本的な場合について典型的な結果を得るのを目標とすべく問題の単純化を行う．一般化することは原理的に可能であると思うが計算を複雑にする．次の仮定を置く．
仮定１：受験生をある一つの能力の大小に関して一列に並べることができる．入学試験において合否の判定を入試成績に基づいて行うということは，受験生を入試成績に基づくあるやり方で一列に並べて上から定員分だけ取るということであるから，最終的にはある一つの実数値を以て能力としているはずである．例えば数学と英語を試験として課す場合でも合否判定の際は合計点という形で一つの数字を見ている．合否を決めるために一列に並べるその順序づけを数値で表したものという意味で一つの実数値を考える．
この数値を実際の試験において具体的にどう計算するかは重要な問題である．生の得点を用いるか偏差値等を用いるか，科目間の比重をどう取るかなどの問題については宮沢（前出）３，４章を参照．
能力を実数値 x で表す．
仮定２：受験生の能力は均一ではない．それ故に，能力を測定して多い順に取るという考えが成り立つのであるから，これは試験を行う根底にある仮定と言って良いだろう．選抜試験は能力自体が受験生間で均一ではなく分布を持つことが本質である．受験生の能力 x の normalization （0 の目盛りの位置と単位）を適当にとって x の平均を 0，分散を 1 としておく．
仮定３：試験成績が正確に受験生の能力を測るならば能力順に合格するはずであるが，実際は初めに議論したように各受験生の成績はその人の能力の周りにばらつく．これが即ち試験成績の誤差である．この誤差は試験のやり方，時間（問題量）などによって変わり得る量である．裏返して言えば，この誤差を適切に調節することにより与えられた制約の下で最も合理的な試験方法を選択する自由度が，試験をする側に残っているのである．

以上の仮定は，入学試験を能力の測定とする立場からは本質的な仮定と思われる．以下の仮定は議論および計算を簡単にするためのものである．
仮定４：問題にしている能力以外の個人差を無視する．能力 x を持つ各々の受験生について試験成績は平均 x で分布する．この個人の成績分布は x と試験だけで決まり，受験生の能力以外によって片寄ることはない．
仮定５：能力，成績ともに正規分布する場合を考える．能力は仮定２より平均 0，分散 1 の正規分布，能力 x の受験生の成績は平均 x，分散 σ^2 の正規分布とする．仮定３で注意したように試験によって σ が異なる．σ が小さいほど試験の精度が高い．宮沢（前出）では σ の代わりに信頼性係数 r を用いている．σ と r の関係は

$$r = (1 + \sigma^2)^{-1/2} . \qquad (2.1)$$

$0 \leqq r \leqq 1$ であって，$r=0$ が $\sigma=\infty$（不正確），$r=1$ が $\sigma=0$（正確），に対応する．

実測値からσ（即ち信頼性係数r）を決める方法については宮沢（「入学試験の誤差と信頼性」（研究会Ⅳ））を参照．一般に大学入試においては信頼性係数 r は 1 に近い，即ちσは比較的小さい．
仮定６：複数の試験を行うときは試験の間には直接の相関がない．例えば多段階選抜方式において，能力 x の受験生の１次試験と２次試験の成績は互いに独立に x の周りに分布する．これは１次と２次で似た問題を出したり，２次の成績判定で１次の成績を加味する，というケースを簡単のため除いて考えるということである．
注意：１次と２次は同じ能力 x を測っているので，x を積分すれば試験成績間には相関がある．上記の仮定は受験生 x を決めたときの成績分布の独立性である．
仮定７：大学は一つだけであり，母集団としての受験生は変動しない．

以上の仮定の下で，次のような正確度という量を定義してこれを考察の対象とする：

試験の正確度 R = 通るべき受験生のうち実際に通った受験生の割合．

ここで通るべき受験生とは，能力順に定員分合格させたならば通ったであろう受験生，である．R の代わりに不正確度 E = (1－R)/B （B は入学試験の倍率）を用いることもある．E は全受験生を 1 とするとき能力順ならば合格するべきなのに間違って通らなかった受験生の割合である．これは能力順ならば合格し得ないのに合格してしまった割合でもある．
以下で，諸パラメータに対する R の依存性，および試験の諸方式と R との関係を調べる．

３．試験方式と正確度の表式

以下でよく用いるので

$$H(x) = \exp(-x^2/2)\,(2\pi)^{-1/2}\ , \qquad (3.1)$$

$$G(x) = \int_{t>x} H(t)\,dt\ , \qquad (3.2)$$

とおいておく．G(x) は誤差関数である．

Ａ．１回試験

１回の試験の成績得点によって入学者を決める選抜試験を行う場合を考える．試験の倍率を B (> 1) とおく．能力と成績を表す（確率）変数をそれぞれ X，Y と書く．(X,Y) = (x,y) となる点での分布密度を ρ(x,y) とおく．２章の仮定により，ρとして

$$\rho(x,y) = H(x)\ H((y-x)/\sigma)\ /\sigma\ , \qquad (3.3)$$

ととる．
合格最低点（成績）を y=C とする．C は

$$\mathrm{Prob}(\ Y \geqq C\ ) = 1/B\ , \qquad (3.4)$$

で定義される．ここで Prob(…) は … が起こる割合を表す：

$$\mathrm{Prob}(\ Y \geqq C\ ) = \int_{y \geqq C} \rho(x,y)\ dx\ dy\ .$$

試験全体の正確度 R は２章の定義に基づいて

$$R = \mathrm{Prob}(\ Y \geqq C \cap X \geqq d\ )\ /\ \mathrm{Prob}(\ X \geqq d\ )\ , \qquad (3.5)$$

と書ける．ここで d は

$$\mathrm{Prob}(\ X \geqq d\ ) = 1/B\ , \qquad (3.6)$$

で定義される．即ち，本来能力順に数えたときに通るべき人の能力下限である．

(3.4),(3.5),(3.6)は(3.3)を代入して適当な変数変換をすると次のように書き換えることができる．

$$G(C\ (1+\sigma^2)^{-1/2}) = 1/B\ , \tag{3.7}$$

$$E = (1-R)/B$$

$$= \int H(x)\ G\left(\frac{C-x}{\sigma}\right)\ \theta(d-x)\ dx\ , \tag{3.8}$$

θ(x)は $x \geqq 0$ のとき 1, $x<0$ のとき 0 となる関数,

$$G(d) = 1/B\ . \tag{3.9}$$

(3.7),(3.8),(3.9)が諸量の間の関係式である．

Ｂ．多段階選抜方式（足切り方式）

1次試験からN次試験まである足切り型試験($N \geqq 1$)による入学者選抜試験を行う場合を考える．

k次試験($k=1,...,N$)の倍率を B_k (>1) とおく．全倍率 $B = B_1...B_N$.

能力，成績(1)，…，成績(N)を表す確率変数をそれぞれ $X,Y_1,...,Y_N$ と書く．

$(X,Y_1,...,Y_N) = (x,y_1,...,y_N)$ となる点での分布密度を $\rho(x,y_1,...,y_N)$ とおく．1回試験の場合と同様に，ρとして

$$\rho(x,y) = H(x)\ \rho_1(y_1,x)\ ...\ \rho_N(y_N,x)\ , \tag{3.10}$$

$$\rho_k(y_k,x) = H((y_k-x)/\sigma_k)\ /\sigma_k\ ,\quad k=1,2,...,N,$$

ととる．

k次試験に於ける足切り点（合格最低成績）を C_k とする($k=1,..,N$).

事象 $D_k = \{\ Y_k \geqq C_k\ \}$, $k=1,...,N$, とおくと, C_k は

$$\mathrm{Prob}(\ D_1 \cap ... \cap D_k\ ) = 1/(B_1...B_k)\ ,\ k=1,...,N\ , \tag{3.11}$$

で帰納的に定義される．k=N における(3.11)は合格率(= 1/全倍率)を表す．

試験全体の正確度 R は

$$R = \mathrm{Prob}(\ D_1 \cap ... \cap D_N \cap \{X \geqq d\}\ )\ /\ \mathrm{Prob}(X \geqq d)\ , \tag{3.12}$$

で定義される．d は前と同様に(3.6)あるいは(3.9)で定義される．

$E = (1-R)/B = 1/B - R/B$ で書けば

$$E = \mathrm{Prob}(\ D_1 \cap ... \cap D_N \cap \{X<d\}\ )\ . \tag{3.13}$$

N=2 のとき(3.11),(3.13)は次のようになる．

$$\frac{1}{B_1} = \text{1次試験の競争率} = \int H(x)\ G\left(\frac{C_1-x}{\sigma_1}\right)\ dx\ , \tag{3.14}$$

$$\frac{1}{B} = \text{合格率} = \int H(x)\ G\left(\frac{C_1-x}{\sigma_1}\right)\ G\left(\frac{C_2-x}{\sigma_2}\right)\ dx\ , \tag{3.15}$$

$$E = (1-R)/B$$

$$= \int H(x)\ G\left(\frac{C_1-x}{\sigma_1}\right)\ G\left(\frac{C_2-x}{\sigma_2}\right)\ \theta(d-x)\ dx\ . \tag{3.16}$$

C_1(C_2) が１次（２次）試験の合格最低点を表す．

σ_1(σ_2) は１次（２次）試験単独の誤差を表す．

(3.9),(3.14),(3.15),(3.16)が諸量の間の関係式である．

C．分離分割方式

前後期２回試験を行う分離分割方式の入学試験を考える．母集団は２回の試験で共通とする．従って，前期試験で通った分だけ後期試験の受験生が減るとする．前期のみ，後期のみの受験は考えない．今まで同様次の諸量を定義する．

B: 倍率＝全受験生／最終的な合格者総数， B_1: 前期倍率，

C_1 (C_2): 前期（後期）試験の合格最低点，

E=(1－R)/B, R: 正確度．

今までと同様の考察により次の諸式を得る．

$$G(d) = 1/B \;, \qquad (3.9)$$

$$\frac{1}{B_1} = \int H(x)\; G\!\left(\frac{C_1 - x}{\sigma_1}\right) dx \;, \qquad (3.17)$$

$$\frac{1}{B} = \int H(x)\left[G\!\left(\frac{C_1 - x}{\sigma_1}\right) + \left\{1 - G\!\left(\frac{C_1 - x}{\sigma_1}\right)\right\} G\!\left(\frac{C_2 - x}{\sigma_2}\right)\right] dx$$

$$= 1 - \int H(x)\; G\!\left(\frac{-C_1 - x}{\sigma_1}\right) G\!\left(\frac{-C_2 - x}{\sigma_2}\right) dx \;, \qquad (3.18)$$

合格者は前期の合格者と前期の失格者のうち後期の合格者との和であることから(3.18)が導かれる．

$$E = \int H(x)\left[G\!\left(\frac{C_1 - x}{\sigma_1}\right) + \left\{1 - G\!\left(\frac{C_1 - x}{\sigma_1}\right)\right\} G\!\left(\frac{C_2 - x}{\sigma_2}\right)\right] \theta(d - x)\, dx$$

$$= \int H(x)\; G\!\left(\frac{-C_1 - x}{\sigma_1}\right) G\!\left(\frac{-C_2 - x}{\sigma_2}\right) \theta(-d - x)\, dx \;. \qquad (3.19)$$

(3.9),(3.17),(3.18),(3.19)が諸量の間の関係式である．

$$1/B' = 1 - 1/B \;,\quad 1/B_1{}' = 1 - 1/B_1,$$

$$d' = -d \;,\quad C_1{}' = -C_1 \;,\quad C_2{}' = -C_2 \;, \qquad (3.20)$$

とおくと，(3.9)(3.17)(3.18)(3.19)は

$$1/B' = G(d') \;, \qquad (3.21)$$

$$\frac{1}{B_1{}'} = \int H(x)\; G\!\left(\frac{C_1{}' - x}{\sigma_1}\right) dx \;, \qquad (3.22)$$

$$\frac{1}{B'} = \int H(x)\; G\!\left(\frac{C_1{}' - x}{\sigma_1}\right) G\!\left(\frac{C_2{}' - x}{\sigma_2}\right) dx \;, \qquad (3.23)$$

$$E = \int H(x)\; G\!\left(\frac{C_1{}' - x}{\sigma_1}\right) G\!\left(\frac{C_2{}' - x}{\sigma_2}\right) \theta(d' - x)\, dx \;. \qquad (3.24)$$

以上を２段階選抜方式の諸式と比べると，分離分割方式の不正確度 E は変換(3.20)を施したときの２段階選抜方式の不正確度 E に一致することが分かる．この結果は次のように説明できる：

<u>分離分割方式は不合格者を２段階選抜（足切り）方式で選ぶ方式である．</u>

不正確度は間違って合格した受験生の（全受験生に対する）割合とみても良いし，間違って不合格になった受験生の割合とみても良い．よって，分離分割方式の正

確度は足切り方式の正確度の表式で表すことができるということである．

４．正確度の計算

Ａ．１回試験

計算すべき積分は一般には誤差関数までの特殊関数で表すことができないが，試験の精度がよいとき即ちσが 0 に近いとき（信頼性係数 r が 1 に近いとき）はσについて展開することにより，誤差関数で結果を書くことができる．大学入試の場合には比較的試験の精度がよいと考えられる．以下そのような場合の表式を求める．

$$C = d\ (1 + \sigma^2)^{1/2} . \qquad (4.1)$$

σの3次までで

$$E = \frac{\sigma}{2\pi}\ \exp(-d^2/2)\ \{1 - \frac{\sigma^2}{3}\ (1+ \frac{d^2}{8})\}\ , \qquad (4.2)$$

或は，

$$R = 1 - \frac{B\sigma}{2\pi}\ \exp(-d^2/2)\ \{1 - \frac{\sigma^2}{3}\ (1+ \frac{d^2}{8})\}\ . \qquad (4.3)$$

d は

$$G(d) = 1/B\ , \qquad (3.9)$$

から数値積分または正規分布の数表を用いて決める．

σについての展開計算結果(4.2)と元の定義(3.8)を数値積分した結果を比較してみると，相対誤差は倍率$B \leqq 5$ において，信頼性係数r=0.8 のとき 6%程度，r=0.9 のとき 1%以下である．従って，この領域では展開計算は十分な精度を持っている．

表：試験の精度σ（信頼性係数 r ）を決めたときの倍率Bと正確度Rの関係

r= .8000 (σ^2= .5625)				r= .9000 (σ^2= .2346)			
B	R	c	d	B	R	c	d
1.2	.93	-1.3	-1.0	1.2	.95	-1.1	-1.0
1.4	.89	-0.8	-0.6	1.4	.92	-0.7	-0.6
1.7	.83	-0.3	-0.2	1.7	.88	-0.2	-0.2
2.0	.80	0.0	0.0	2.0	.86	0.0	0.0
2.4	.76	0.3	0.2	2.4	.83	0.2	0.2
3.6	.69	0.8	0.6	3.6	.78	0.7	0.6
6.3	.62	1.3	1.0	6.3	.73	1.1	1.0

Ｂ．２段階選抜方式

１次（２次）試験の倍率を B_1 (B_2) とする ($B=B_1B_2$)．

これらが与えられていると，

d: 能力順でみたとき最終合格者となるべき能力下限，

C_1 (C_2): １次（２次）試験の合格最低点，

$R = 1-EB$: 正確度，

はそれぞれ(3.9)(3.14)(3.15)(3.16) によって求められる．

実際の試験では全倍率 B は与えられているが１次試験の倍率B_1は自由に決めることができるから，２章の議論によれば正確度 R を最大にするように B_1を定めるべきである．現実の２段階選抜方式入学試験においては１次試験の倍率は経験則や慣習で決められている場合があるのではないだろうか．我々は１次試験の倍率を定める合理的方法として，入学試験全体の正確度 R を最大にするよう決めるという基準を提案する．

$\sigma_1=\sigma_2$，即ち１次試験と２次試験の精度（信頼性係数）が等しい場合には正確度 R を最大にする１次試験の倍率 B_1 は次のような考察で求められる．全倍率 B が与えられているとしよう．(3.9)により d が定まる．$\sigma_1=\sigma_2$ の場合，(3.16)から正確度 R は C_1 と C_2 の入れ替えに関して不変であることが分かる．よって R が最大になるのは $C_1=C_2$ となるように（各次試験の倍率を）選んだときである．

<u>σが１次試験と２次試験で等しい場合， R が最大になるのは</u>
<u>C_1とC_2が等しくなるように各次試験の倍率を選んだときである．</u>

このようにとった場合を考える．各次試験の倍率と全体の正確度を展開計算で求めると次のようになる（σの3次までで）：

$$G(d) = 1/B ,$$

$$C = d - \sigma/\sqrt{\pi} + \frac{d}{2}\left(1-\frac{1}{\pi}\right)\sigma^2 + (d^2-1)\left(1-\frac{4}{\pi}\right)\sigma^3/(12\sqrt{\pi}) ,$$

$$1/B_1 = G(C) + \sigma^2 CH(C)/2 ,$$

$$B_2 = B/B_1 ,$$

$v = (C-d)/\sigma$ とおくと

$$E = \sigma H(C)\{B + G(v)H(v) + C\sigma(H(v)^2 + 2A)/2 + (C^2-1)\sigma^2(5B + 3G(v)H(v) + 4vA + 2vH(v)^2)/12\} ,$$

$$A = vG(v)H(v) - H(v)^2 ,$$

$$B = \pi^{-1/2}\{G(v)^2 - G(2^{1/2}v)\} + H(v)G(v) ,$$

$$R = 1 - E B .$$

注目すべきことは，$\sigma\to 0$ のとき $B_2\to 1$ 即ち試験の精度がよい場合には

<u>１次試験による足切りは相当大きくとるのがよい．</u>

一見奇妙であるが，１次試験と２次試験が同じ精度（例えば試験時間が同じ程度）であるにも関わらず，１次試験で殆ど合否を決めてしまい２次試験の倍率は１に近くとる（下表参照．全倍率が 10 倍でも２次試験の倍率は 2 以下であることに注意）のが全体としてむしろ正確であるということである．正確度の計算に基づく発見の一つと言えよう．

表：１次試験と２次試験の精度σ(r)が等しい場合，適切な１次試験倍率B_1の設定とそのときの正確度R．例：各次試験が信頼性係数r=0.9の場合，全倍率が2倍ならば１次試験の倍率は1.7，２次試験の倍率は1.2とするのが正確度最大，全倍率が10倍ならば２次試験の倍率は1.6がよい．

r=0.80　σ^2=0.56	r=0.90　σ^2=0.23	r=0.95　σ^2=0.11
B　B_1　B_2　R	B　B_1　B_2　R	B　B_1　B_2　R

2	1.6	1.2	.83	2	1.7	1.2	.88	2	1.8	1.1	.92
3	2.2	1.4	.77	3	2.3	1.3	.84	3	2.5	1.2	.89
4	2.7	1.5	.73	4	2.9	1.4	.81	4	3.2	1.3	.87
5	3.1	1.6	.71	5	3.5	1.4	.79	5	3.9	1.3	.85
6	3.5	1.7	.69	6	4.1	1.5	.78	6	4.6	1.3	.84
8	4.3	1.8	.66	8	5.2	1.5	.76	8	5.9	1.4	.83
10	5.1	2.0	.64	10	6.2	1.6	.74	10	7.2	1.4	.83

$\sigma_1 \neq \sigma_2$, 即ち1次試験と2次試験の信頼性が異なる場合にも(3.9),(3.14),(3.15),(3.16)を元にして正確度 R が最大になるように1次試験の倍率B_1を定めれば良い．$\sigma_1=\sigma_2$ の場合に比べると煩雑な結果になるが，$\sigma_1=\sigma_2$ の場合からのずれが小さい場合について，計算のプロセスは省略して結果を述べる．

$$\sigma^{-1} = \{\sigma_1^{-1} + \sigma_2^{-1}\} / 2 ,$$

$$\varepsilon = \{\sigma_1 - \sigma_2\} / \{\sigma_1 + \sigma_2\} ,$$

とおき，σとεを小さいとしてこれらについてべき展開したときにこれらの合計3次までで次のようになる．

$$B = 1/G(d) .$$

$$C_1 = C + \sigma_1 \varepsilon \xi_1 ,$$

$$C_2 = C - \sigma_2 \varepsilon \xi_2 ,$$

$$C = d - \sigma/\sqrt{\pi} + \frac{d}{2}\left(1-\frac{1}{\pi}\right)\sigma^2 + (d^2-1)\left(1-\frac{4}{\pi}\right)\sigma^3/(12\sqrt{\pi}) ,$$

$$\frac{\xi_2 - \xi_1}{\varepsilon} = -3v + 2A/B - 2vA^2/B^2 ,$$

$$\frac{\xi_1 + \xi_2}{2} = \{A + C\sigma(B+2vA)\} / \{B - C\sigma H(v)^2\} ,$$

$$v = \frac{C-d}{\sigma} ,$$

$$A = vG(v)H(v) - H(v)^2 ,$$

$$B = \pi^{-1/2}\{G(v)^2 - G(2^{1/2}v)\} + H(v)G(v) .$$

$$1/B_1 = 1/B_1{}^{\cdot} - \sigma\varepsilon(1+\varepsilon)\xi_1 H(C) + \sigma^2\varepsilon C H(C) ,$$

$$1/B_1{}^{\cdot} = 1/B + \sigma H(C)\{1 + \frac{5}{12}\sigma^2(C^2-1)\} / \sqrt{\pi} .$$

$$R = 1 - EB ,$$

$$E = E^{\cdot} + \frac{1}{2}\sigma\varepsilon^2 H(C)\{3B + H(v)G(v) + 2vA - 2A^2/B\} ,$$

$$E^{\cdot} = \sigma H(C)\{B + G(v)H(v) + C\sigma(H(v)^2 + 2A)/2 + (C^2-1)\sigma^2(5B + 3G(v)H(v) + 4vA + 2vH(v)^2)/12\} .$$

C．分離分割方式（$\sigma_1=\sigma_2$の場合）

(3.20)の対応と2段階選抜方式の場合の計算結果により分離分割方式の入試の正確度を計算することができる．ここでは，$\sigma_1=\sigma_2$の場合（前期試験と後期試験の精度が等しい場合）についてσの2次までで分離分割方式と2段階選抜方式の正確度 R を比較してみる．結果は，

$$R[選抜] = 1-\sigma\ B\ H(d)\ p + \sigma^2\ B\ H(d)\ d\ q\ ,$$
$$R[分割] = 1-\sigma\ B\ H(d)\ p - \sigma^2\ B\ H(d)\ d\ q\ .$$

ここで p, q は σ にも B にもよらない定数で $r=-1/\sqrt{\pi}$ とおくとき

$$p = -r\ G(r)^2 + 2\ H(r)\ G(r) + r\ G(r\sqrt{2}) \quad 0.3287\ ,$$
$$q = -r^2 G(r)^2 + r\ H(r)\ G(r) + r^2 G(r\sqrt{2}) + H(r)^2/2 \quad 0.0094\ .$$

d は(3.9)によって定まるが，$B>2$ のとき正で $B<2$ のとき負になるので，

$B > 2$ のとき $R[選抜] > R[分割]$,

$B < 2$ のとき $R[選抜] < R[分割]$,

となる．同程度の２回分の試験問題が２つあるときに，これを２段階選抜方式として利用するか分離分割方式として利用するかを判断できる．

倍率が 2 より大きいときは選抜方式の方が分離分割方式より正確である．

５．試験時間の配分

４章では１次・２次試験の時間や問題が各々定まっているとして正確度を計算した．従って特に各次試験の精度 σ_i（信頼性係数 r_i）は定まっていた．実際には１次・２次試験の時間配分を変えることにより各次試験の精度も動かせる．

１次試験で落ちた受験生に対しては２次試験を行わなくてよいので労力は節約される．ここで労力とは受験生が受験のために消費する延べ労力あるいは延べ時間のことである．多段階選抜方式でも限りなく試験を行えばいくらでも正確度は増大する（６章参照）が，時間的にも労力の有効利用の観点からもこれは現実的でもないし，受験生の労力の浪費という観点から望ましいことでもあるまい．従って，労力一定の条件下で正確度が最大になるように１次・２次試験の時間を配分する方法を探すのは現実的な課題である．

以下，労力を試験時間で測ることにする．同程度の試験を二日やれば一日分の２倍の労力ということである．適当に時間及び労力の単位を取れば，延べ労力が一定（=1 とおく）という条件は次のように書ける．

１回試験の場合，試験時間 t は，

$$t = 1\ . \tag{5.1}$$

２段階選抜方式の場合，１次（２次）試験の時間を t_1（t_2）とすると，

$$t_1 + t_2/B_1 = 1\ . \tag{5.2}$$

ここで B_1 は1次試験の倍率．

分離分割方式の場合，前期（後期）試験の時間を t_1（t_2）とすると，

$$t_1 + t_2(1-1/B_1) = 1\ . \tag{5.3}$$

ここで B_1 は前期試験の倍率．

他方，時間 t の試験の精度をどう見積るか問題がある．ここでは，能力の周りの成績の分布の分散を

$$\sigma^2 = \sigma_0^2/t\ , \qquad \sigma_0 \text{ は定数}, \tag{5.4}$$

と評価することにする．一日４時間より二日８時間の試験の方が σ^2 が半分になるという考えである．

試験が独立，同等な多数の小問からなっていて時間と解答し得る小問数が比例するときは(5.4)が成立するであろう．時間の代わりに出題問題数を用いることも考えられるが，一問の単位が明確に定義できないという問題点がある．単純な計算問題ならばほぼ独立・同等な多数の問題を用意できるかも知れないが，大学入学試験問題ではそれは難しいであろう．

注意：(5.4)にも全く問題がないわけではない．試験は時間をかけて十分問題を解かせるほど能力評価が正確になるというのは正しいであろうが，試験問題が決まっていると十分時間が立てば飽和してしまう．つまり，個々の学生にとってある程度以上時間を延ばしても答案改善の余地がなくなれば精度が時間と共に伸びなくなってしまう恐れがある．ここではこの問題には深入りせず，(5.4)を採用して話を進める．

以上の条件の下で正確度を大きくしたい．１回試験については４章Ａの結果に(5.1),(5.4)を代入すれば直ちに結果を得る．２段階選抜方式については(3.14),(3.15),(3.9),(5.2),(5.4)を用いて計算すればよい．分離分割方式の問題は変換(3.20)によって２段階選抜方式の問題に帰着する．表に各々の場合について近似計算を行って最適解を探した結果をまとめておく．

表：２段階選抜方式，労力一定の条件で適切な１次・２次試験の時間配分比 t_2/t_1 及び１次（２次）試験の倍率 B_1,B_2 とその時の正確度R[選]．r_1 (r_2) は１次（２次）試験の信頼性係数．R[１]，R[分]は同じ延べ労力でそれぞれ１回試験及び適切に前後期の比率をとった場合の分離分割方式試験の正確度．

$r_0=0.90$	$\sigma_0^2=0.23$							
B	t_2/t_1	B_1	B_2	r_1	r_2	R[選]	R[１]	R[分]
2	1.2	1.6	1.3	0.84	0.86	0.85	0.85	0.85
3	1.4	2.3	1.3	0.85	0.89	0.81	0.79	0.78
4	1.6	2.9	1.4	0.86	0.90	0.79	0.75	0.73
5	1.9	3.6	1.4	0.86	0.92	0.78	0.73	0.70
6	2.0	4.3	1.4	0.86	0.93	0.77	0.71	0.68
8	2.3	5.6	1.4	0.87	0.93	0.75	0.68	0.64
$r_0=0.95$	$\sigma_0^2=0.11$							
B	t_2/t_1	B_1	B_2	r_1	r_2	R[選]	R[１]	R[分]
2	1.3	1.7	1.2	0.92	0.93	0.90	0.90	0.90
3	1.6	2.5	1.2	0.92	0.95	0.87	0.86	0.85
4	1.8	3.2	1.3	0.92	0.96	0.86	0.83	0.82
5	2.0	4.0	1.3	0.93	0.96	0.85	0.82	0.80
6	2.2	4.7	1.3	0.93	0.97	0.84	0.80	0.78
8	2.5	6.2	1.3	0.93	0.97	0.83	0.78	0.76

競争率が2より大きい場合，表より以下のことが読み取れる．

(1)　労力（受験生の延べ受験時間）を一定にとるならば，１回の試験で入学者を決めるより２段階選抜（足切り）方式で決めた方が正確である．判定のための労力を足切りによって境界付近の受験生に集中するからである．正確度という判断基準からは足切りは正しい考え方と言える．分離分割方式については，前期（後期）試験の倍率や試験量の配分比をどう選んでも同じ労力の２段階選抜方式より正確度が落ちる．全倍率が大きいほど２段階選抜方式と他の方式（１回試験や分離分割方式）の正確度の差は大きくなり，その効果は顕著である．

(2)　最適の足切り倍率や１次・２次試験の試験量の比は倍率や全試験量に依存するが，足切り倍率はかなり大きくとるのが全体の正確度を高くする．最終的な倍率が2-8倍で各次試験の精度が比較的良い場合の計算結果によれば，２次試験の倍率は2倍を越えない．全倍率が大きくなっても２次試験の倍率は極めてゆっくりとしか大きくならない．この結論は試験全体の精度が良いほど顕著であって，試験の精度が高くなると２次試験の倍率は1に近く（１次試験の倍率は全倍率に近く）

とるのが適切になる．
(3)　１次・２次試験の試験量の比は２次試験の方に重点をおくのがよい．即ち２次試験の方の時間を長くとって精度をよくする方が正確度が高い．この傾向は全倍率が大きいほど，試験の精度が高いほど顕著である．全倍率が小さいときは１次試験と２次試験の試験時間（試験量）はほぼ等しくとるのがよいが，全倍率が上がるにつれて２次試験の相対的な量が増える．倍率が高くなると足切りによる労力の節約が大きくなるので，１次の精度をわずかに下げても２次の精度を大きく改善できることによって全体としての正確度が向上するからである．試験の精度によるが，全倍率が5,6倍を越えると２次試験と１次試験の量の比が2を越える．
　競争率が2より小さい場合は2より大きい場合とは逆に，分離分割方式が最善となる．これは，分離分割方式と２段階選抜方式の対応関係（３章Ｃ参照）から理解できる．対応関係(3.20)により，
$1/B \to 1-1/B$，$1/B_1 \to 1-1/B_1$，$1/B_2 \to 1-1/B_2$，
と読み替えれば２段階選抜方式の結果をそのまま利用できる．上記表の B，B_1，B_2 の欄を読み替えれば t_2/t_1，r_1，r_2 はそのまま使える．（このとき不変なのは不正確度 $E=(1-R)/B$ なので，R は読み替えを要することに注意．）
　競争率がちょうど2のときは（各次試験の倍率及び試験時間配分を適切に選んだ）２段階選抜方式と分離分割方式は正確度が等しく，１回試験の正確度に優るが，その差は僅かである．競争率２のとき諸変数は下表の選択が最善となる．

		足切り方式		分離分割方式	
r_0	t_2/t_1	B_1	B_2	B_1	B_2
0.9	1.2	1.6	1.3	2.7	4.8
0.95	1.3	1.7	1.2	2.4	6.6

６．極限定理

　多段階選抜方式の正確度に関して次のような数学的問題を考える：多段階選抜方式の場合，段階数Nを増やせば正確度は1に近づくだろうか，それとも多段階選抜方式は原理的な不正確があって，段階数を変えても正確度は1に近づかないだろうか？これについては次の数学的事実が証明できる：全倍率 B (>1)を固定する．信頼性係数 r（誤差 σ）と足切り点 $C(N)$ を共通にとったN段階選抜試験を考える．この時 N を大きくしていくと正確度 R は 1 に近づく（$E=(1-R)/B$ は 0 に近づく）．その近づき方は，

$$E(N) = b\sigma H(d) \,/\, x(N) + \text{高次項（速く消える項）}, \qquad (6.1)$$

で与えられる．ここで定数 b は，

$$b = \int_{-\gamma}^{\infty} \exp\{-\exp(x)\}\, dx ,$$

$$\gamma = -\int_{0}^{\infty} \exp(-y)\log y\, dy \quad (\text{Eulerの定数}),$$

で定義され，$x(N)$ は

$$x(N)/(N-1) = H(x(N))/(1-G(x(N))) \qquad (6.2)$$

を満たす実数．関数 H, G は(3.1),(3.2)で定義され，d は全倍率 B を用いて(3.9)によって定義されている．$x(N)$ は(6.2)でただ一つ決まる正数で，Nが大きいときの漸近的な振舞いは

$$x(N) = (2 \log N) \quad + (N\to\infty\text{で消える項}) , \qquad (6.3)$$

で与えられる．

(6.1),(6.3)より $E(N)\to 0$ であるから多段階選抜方式には原理的な限界がないことが分かる．試験の回数 N を大きくすると E(N) が log N の平方根に反比例して小さくなることも分かるが，これは中心極限定理（N 回の試験の平均点で合否を決定する場合に対応）と比べると大変遅い．中心極限定理では N の平方根に反比例して誤差が小さくなることが知られている．なお，問題で足切り点 C を共通にとったのは，４章Ｂによって r が共通の時は足切り点を揃えるのがよいことが分かるからである．

以上の結果は(3.9),(3.11),(3.13)より得られる

$$1/B = \int_{x>d} \exp(-x^2/2)\, dx\, (2\pi)^{-1/2} ,$$

$$1/B = \int \exp(-x^2/2)\, [G(\{C(N)-x\}/\sigma)]^N\, dx\, (2\pi)^{-1/2} ,$$

$$E(N) = \int_{x<d} \exp(-x^2/2)\, [G(\{C(N)-x\}/\sigma)]^N\, dx\, (2\pi)^{-1/2} ,$$

及び，直接数学的に証明できる

$$\lim_{N\to\infty} G(-x(N)-x/x(N))^N = \exp\{-\exp(-x)\} ,$$

を用いて導かれる．なお，足切り点 C(N) の漸近形は

$$C(N)^2 = (\sigma x(N)-d)^2+2\gamma\sigma^2 + o(N^0) , \quad C(N) < 0 ,$$

となる．$N\to\infty$ で $C(N)\to-\infty$ となるのは一つ一つの試験では少しずつ裾野だけを落として N 回全体で最初に設定した倍率に至るからである．