欠席率を考慮した第１段階選抜倍率の決定について

千葉大学　田栗　正章，橋本　明浩，今野　良彦

1　はじめに

大学入試センター試験を利用する大学入試において，センター試験の成績により第１段階選抜を行う場合がある。その理由としては，次のようなものが考えられる。

(A)　個別学力試験で実験や実習を行わせる必要があるが，実験器具等の数の制約のため第１段階選抜を行う。また個別学力試験で面接を行う必要があるが，時間等の制約から第１段階選抜を行う。

(B)　大学入試センター試験と個別学力試験の合計で最終的な合否を決定する場合，個別学力試験での逆転が，ほぼ不可能と考えられるため，第１段階選抜を行う。

ところで，第１段階選抜は可能な限り行わない方が望ましいと考えられる。実際，個別学力試験で満点をとっても最終的に逆転合格が不可能な場合でさえ，受験料を返却することが困難な状況では，第１段階選抜は行うべきでないとの議論もある。また今年度のように，大学入試センター試験の平均点に科目間格差がある等の理由により，できるだけ第１段階選抜を行わないようにとの指導が行われる場合もある。

そこで本論では，上記(A)のような場合に，第１段階選抜で不合格となる志願者数をできるだけ少なくするための１つの試みについて考察する。上記(B)の問題については，矢野他(1990)に１つの考え方とその解法が提案されている。

表１には，千葉大学の４つの学科の過去５年間の後期試験における志願者数，受験予定者数(第１段階選抜後の人数)，個別学力試験(以下２次試験と呼ぶ)の受験者数が示されている。この表の詳しい説明は，2.1節に与えてある。これより，学科にもよるが例年かなり（３～７割程度）の欠席者のいることが分かる。例えば平成５年度のＡ学科では，第１段階選抜を行って102名の志願者を35名にしている。ところが実際に個別学力試験を受験した人数は13名であり，欠席者数は22名（欠席率は62.9%）であった。したがって結果的にいえば，第１段階選抜では35名にまで減らす必要はなく，この観点だけからは，少なくとももう22名は第１段階選抜で合格としてもよかったことになる。このような事情は，程度の差こそあれ，多くの大学の選抜単位で起こりうるものであろう。

それでは個別学力試験における欠席者数をどの位に見込んで，第１段階選抜合格者数を決定すればよいのであろうか。この問題は，欠席者を見込んで第１段階選抜の合格者を実験器具や時間等の制約を越えてとった場合，見込んだほど欠席者がいなかったときに生じるリスクがどの程度重大であるかに依存して，解析の方法を考えるべきであろう。しかしここではこの問題を，欠席者数が各種の制約を越える確率が十分小さくなるように受験予定者数を決定するという問題として捉え，それに対する１つの解法を検討する。次の第２節では，この解析のためのモデルを作成し，それに基づく解析方法を与える。実際データを用いての解析結果とその検討は，第３節で議論する。

最後の第４節では，得られた結果をふまえて，いくつかの留意点を述べる。

表1：志願者と2次試験受験状況

年度	A学科			B学科			C学科			D学科		
	志願者（倍率）	受験予定者（倍率）	受験者［受験率］	志願者（倍率）	受験予定者（倍率）	受験者［受験率］	志願者（倍率）	受験予定者（倍率）	受験者［受験率］	志願者（倍率）	受験予定者（倍率）	受験者［受験率］
5	102 (25.3)	35 (8.8)	13 [37.1]	88 (5.9)	51 (3.4)	37 [72.5]	182 (10.7)	120 (7.1)	78 [65.0]	55 (9.2)	42 (7.0)	27 [64.3]
6	76 (15.2)	40 (8.0)	22 [55.0]	128 (8.5)	61 (4.1)	27 [44.3]	153 (9.0)	119 (7.0)	80 [67.2]	47 (7.8)	52 (7.0)	29 [69.0]
7	99 (19.8)	40 (8.0)	13 [32.5]	99 (6.6)	65 (4.3)	37 [56.9]	173 (10.2)	119 (7.0)	79 [66.4]	48 (8.0)	48 (8.0)	27 [56.3]
8	133 (16.6)	72 (9.0)	36 [50.0]	102 (6.8)	63 (4.2)	34 [54.0]	172 (10.1)	119 (7.0)	77 [64.7]	53 (8.8)	42 (7.0)	26 [61.9]
9	144 (10.3)	143 (10.2)	89 [62.2]	241 (8.0)	150 (5.0)	77 [51.3]	177 (6.8)	177 (6.8)	121 [68.4]	55 (9.2)	55 (9.2)	29 [52.7]

2　欠席者を考慮した受験予定者数の決定

2.1　与えられるデータと問題の定式化

表1には，千葉大学の4つの学科（A，B，C，D学科とした）の，平成5年度から9年度にわたる過去5年間の後期試験における志願者数，受験予定者数（第1段階選抜後の人数），個別学力試験（2次試験）の受験者数を示している。各マス下段の（　）内の数字は定員に対する倍率を，［　］内の数字は受験予定者数に対する受験者数のパーセント（受験率）を表している。ここで，平成9年度入試においては，A，C，D学科では第1段階選抜を行わなかった。平成9年度のA学科で，志願者数と受験予定者数が食い違っているのは，指定された選択科目の誤りによる失格者が1名いたためである。

この表中の［　］内に示された受験率を見ると，C，D学科ではその値は年度によらず比較的安定しているが，A，B学科の値は年度によるバラツキが大きい。したがって解析に際しては，この点を考慮に入れるべきであろう。

いま，機器や時間等の制約から2次試験を受けさせることのできる最大人数を a，第1段階選抜合格者数（受験予定者数）を n，2次試験受験者数を X とし，$R=X/n$（受験率）とおく。このとき問題は，欠席者数が少なくて各種の制約を越えてしまう確率が十分小さくなるように受験予定者数を決定することである。言い換えれば，受験者数 X が制約限度人数 a 以下である確率 $Pr\,[R\leq a/n]$ を十分大きく（例えば $1-\alpha$ 以上に）するように受験予定者数 n を決定することである。すなわち

$$Pr\ [R\leq a/n]\ \geq 1-\alpha$$

が成り立つように n を決定することが問題である。

2.2　受験率が安定している場合の解析

第1段階選抜合格者の受験行動について最も単純なモデルは，各受験生が一定の確率 p で2次試験を受験し，$1-p$ で欠席するというものであろう。このようなモデルの下では，

2次試験受験者数 X は2項分布にしたがう確率変数となり，したがって2.1節で定義した R も確率変数となる。このとき R の標本分布は，中心極限定理により近似的に正規分布 $N\ (p,p(1-p)/n)$ となる。したがって，次式が成立しなければならない。

$$Pr[\frac{R-p}{\sqrt{p(1-p)/n}} \leq \frac{a/n-p}{\sqrt{p(1-p)/n}}]=1-\alpha.$$

ここで，z_α を標準正規分布の上側100 α パーセント点とすれば，

$$\frac{a/n-p}{\sqrt{p(1-p)/n}}=z_\alpha$$

でなければならない。これを n について解けば，$n>0$ より n は次式を満たさなければならない。

$$n=\frac{4a^2}{\{z_\alpha\sqrt{p(1-p)}+\sqrt{z_\alpha^2 p(1-p)+4ap}\}^2} \qquad (1)$$

したがって，年度によって受験率が安定している場合には，(1)式を用いて第1段階選抜合格者数 n を決定すればよいが，その具体的な手順は以下の通りである。

1° 実験器具の数や面接時間等を考慮して，2次試験を受けさせることのできる最大人数 a を決定する。

2° 実験器具の数の制約や面接時間の制約等がどの程度重要であるのかに応じて，それに対応する z_α の値を標準正規分布表から求める。

3° 過去のデータから，p のおおよその値を求める。例えば過去数年間の受験率の平均値を計算し，それを p の値とする。

4° (1)式を用いて，n の値を計算する。

[註] (1)において近似的な計算を行うと，$4a/z_\alpha^2$ の値が1に比べて十分大きい場合には，p の値が非常に1に近くない限りは，n は p に関して単調減少となっていることが分かる。したがって3°の手順においては，p の値を少し大きめに与えておいた方が安全である。

2.3 受験率が変動する場合の解析

本節では，表1におけるA，B学科のように，年度によって受験率がかなり変動する場合についての検討を行う。3.1節の図1で示すように，特に受験率 p が小さい場合には，(1)によって求めた第1段階選抜合格者数 n の値が大きく変化する。そこで，年度毎に2次試験出席率がかなり変動する場合には p を確率変数と考え，十分大きな確率で p が存在すると考えられる（片側）区間の上端の値を計算し，その値を用いて n を決定することを考える。

いま，ある学科（選抜単位）における過去K年度分の受験率のデータ（n_k，r_k）($k=1,2,\cdots\cdots,K$）が与えられたとする。ここで n_k は第 k 年度の受験予定者数で，$n=\sum_{k=1}^{K}n_k$，$w_k=n_k/n$ とする。また，X_{kj}を第 k 年度の j 番目の受験生が受験するか否かを表す確率変数とし，$R_k=\sum_{j=1}^{nk}X_{kj}/n_k$ とおく。すなわち Rn_k は第 k 年度の受験率を表す確率変数であり，その実現値を r_k とする。ここで，確率変数 X_{kj} および第 k 年度の受験率 p_k（確率変数）の分布について，次の仮定をおく。

① X_{kj} は，互いに独立にパラメータ p_k をもつベルヌイ分布にしたがう。

② p_k は，互いに独立にパラメータ β，γ をもつベータ分布にしたがう。

このとき問題は，考察対象とする学科における第（$K+1$）年度の受験予定者数 n_{K+1} を決定することである。ところで $n_{K+1}=a/p_{K+1}$であればよいから，結局第($K+1$)年度の受験率 p_{K+1}を決定すればよい。したがって，実験器具の数や面接時間等の制約から要請される信頼度を $1-\alpha$ とすれば，次式を満たす受験率 p_{K+1} の上限 c_α を決定すればよいことになる。

$$Pr\ [p_{K+1} \leq c_\alpha] = 1 - \alpha \qquad (2)$$

ところで上の②より，$p_k + 1$ はベータ分布にしたがうと仮定しているが，そのパラメータ β, γ の値は推定しなければならない。与えられたデータからこの分布の平均と分散は推定できるので，ここではモーメント法により，β, γ の推定を行うことにする。ベータ分布の平均，分散を μ，σ^2とし，データから計算される平均，分散を m，s^2とすれば，モーメント推定のための方程式は次式で与えられる。

$$\begin{cases} \mu = \dfrac{\beta}{\beta+\gamma} = m, \\ \sigma^2 = \dfrac{\beta\gamma}{(\beta+\gamma+1)(\beta+\gamma)^2} = s^2. \end{cases}$$

これを β，γ について解けば，次のモーメント推定値 $\hat{\beta}$，$\hat{\gamma}$ を得る。

$$\hat{\beta} = m\hat{\delta}, \qquad \hat{\gamma} = (1-m)\,\hat{\delta}.$$

ただし，$\hat{\delta} = \dfrac{m(1-m)}{s^2} - 1$ である。これらの値を用いれば，(2)式における c_α の推定値 $\hat{c}_\alpha$ を求めることができ，したがって第$(K+1)$年度の受験予定者数 n_{K+1}を決定することが可能となる。具体的な手順は次の通りである。

1° 実験器具の数や面接時間等を考慮して，2次試験を受けさせることのできる最大人数 a を決定する。

2° 実験器具の数の制約や面接時間の制約等がどの程度重要であるのかに応じて，要請する信頼度 $1-\alpha$ を決定する。

3° 与えられたデータから，次式により m，s^2を計算する。

$$m = \sum_{k=1}^{K} w_k r_k,$$

$$s^2 = \frac{1}{1-\sum_{k=1}^{K} w_k^2}\left(\frac{n-1}{n-K}\sum_{k=1}^{K} w_k r_k^2 - m^2 - \frac{n-1}{n-K}m\right).$$

4° 次式により β，γ のモーメント推定値を計算する。

$$\hat{\beta} = m\hat{\delta}, \qquad \hat{\gamma} = (1-m)\hat{\delta},$$
$$\hat{\delta} = \frac{m(1-m)}{s^2} - 1.$$

5° 数値積分法等を適用して，次式を満たす c_α の推定値を求める。

$$\int_0^{c_\alpha} f_p(p;\hat{\beta},\hat{\gamma})\,dp = 1-\alpha,$$

ここで $f_p(p;\hat{\beta},\hat{\gamma})$ はベータ分布の確率密度関数である。

6° 次式により，第$(K+1)$年度の受験予定者数 n_{K+1}の値を計算する。

$$n_{K+1} = \frac{a}{\hat{c}_\alpha}$$

3 受験予定者数の予測結果と実際の結果との比較

3.1 受験率が安定している場合

本節では，2.2節で述べた方法により，考察対象とする学科の受験予定者数（第1段階選抜合格者数）を決定する計算例を与える。

［例1］平成9年度のC学科の場合

表1のC学科では，2次試験に用いる器具の制約から，制約限度人数は $a=120$と考えられた。また，2次試験受験率が予想を越えて高かった場合でも，試験に用いる器具の調達が何とか可能であるような手はずが整えられるので，信頼度は$1-\alpha=0.975$とすることにした。この信頼度に対応する z_α の値は $z_{0.025}=1.96$である。そこでこの場合について，(1)式を用いて p の値を変化させた場合の n の値のグラフを描くと，図1のようになる。ところで表1より，C学科では p の値は年度によらず安定しており，平成5年度から8年度までの受験率の平均値は0.658である。そこで図1から $p=0.66$に対応する n の値を読みとると，$n=160$程度となることが分かる。上で与えた値を(1)式に代入して計算すると，$n=164.3$となる。したがって平成9年度に第

１段階選抜を実施する場合には，$n=164$とすればよさそうであった。

ところで平成９年度の場合には，なるべく第１段階選抜を行わないようにとの指導がなされたので，実際には第１段階選抜は実施されなかった。表１に示されているように，平成９年度の実際の受験率は0.684であった。もし第１段階選抜を行った場合でも，受験率がこのままであったと仮定すれば，実受験者数は$164\times0.684=112.2$となり，制約限度人数120を越えなかったことになる。

ここで2.2節の註で述べたように，少し安全なように p の値を決定した場合を考えてみる。かりに$p=0.70$とした場合には，(1)より$n=155.4$となる。上と同様に考えて受験率が0.684であったとすれば，この場合の実受験者数は$155\times0.684=106.0$となり，制約限度人数120に対し14人程度の余裕があったことになる。

この計算から分かるように，現実の多くの場合においては，p の値を少し安全なように見積もっておいても，第１段階選抜人数の変化はあまり大きくはないと考えられる。

［例２］　平成９年度のＤ学科の場合

表１のＤ学科では，２次試験に用いる器具の制約から，制約限度人数は $a=42$と考えられた。またＣ学科の場合と同様，信頼度は$1-\alpha=0.975$でよいと判断された。表１より，Ｄ学科でも p の値は年度によらず比較的安定しており，平成５年度から８年度までの受験率の平均値は0.629である。そこでこれらの値を(1)式に代入して計算すると，$n=55.6$となる。したがって平成９年度に第１段階選抜を実施する場合には，$n=55$とすればよさそうであった。

ところで平成９年度の場合には，Ｃ学科と同様Ｄ学科においても，実際には第１段階選抜は実施されなかった。表１に示されているように，平成９年度の実際の受験率は0.527であったが，第１段階選抜を行った場合でも受験率がこのままであったと仮定すれば，実受験者数は$55\times0.527=29.0$となり，制約限度人数42を越えなかったことになる。また少し安全なように，p の値を過去の最大値0.690とした場合を考えてみると，(1)より $n=51.4$となる。上と同様に受験率が0.527であったとすれば，この場合の実受験者数は$51\times0.527=26.9$となり，制約限度人数42に対し15人程度の余裕があったことになる。

3.2　受験率が変動する場合

本節では，2.3節で述べた方法により，考察対象とする学科の受験予定者数（第１段階選抜合格者数）を決定する計算例を与え，その計算結果が現実の結果とどの程度符合しているのかを論ずる。

［例３］　平成８年度のＡ，Ｂ学科の場合

表１のＡ，Ｂ学科では，２次試験での面接時間や用いる実験器具の制約から，平成８年度における制約限度人数は，それぞれ $a_A=72$，$a_B=63$と考えられた。また，信頼度については$1-\alpha=0.975$でよいと判断された。また表１から読みとれるように，これらの学科における過去の受験率の値は，かなり大きく変動している。そこでここでは2.3節で与えた手順にしたがって，それぞれの学科に対する n の値を計算してみた。

具体的には，平成５年度から７年度までの過去３年分のデータを用いて，平成８年度における n の値を計算した。この結果は表２にまとめられている。

この表において，推定人数 n は $n=a/\hat{c}_\alpha$ によって計算した。また受験者推定値は，もし第１段階選抜者数を n としたとき，実際に受験したであろうと考えられる人数を意味している。この値は，n に平成８年度におけるそれぞれの学科の受験率の実績値を掛けて得られたものである。例えばＡ学科の場合には，平成８年度の実際の受験率は0.500であったので，受験者推定値は$119\times0.500=59.5$とな

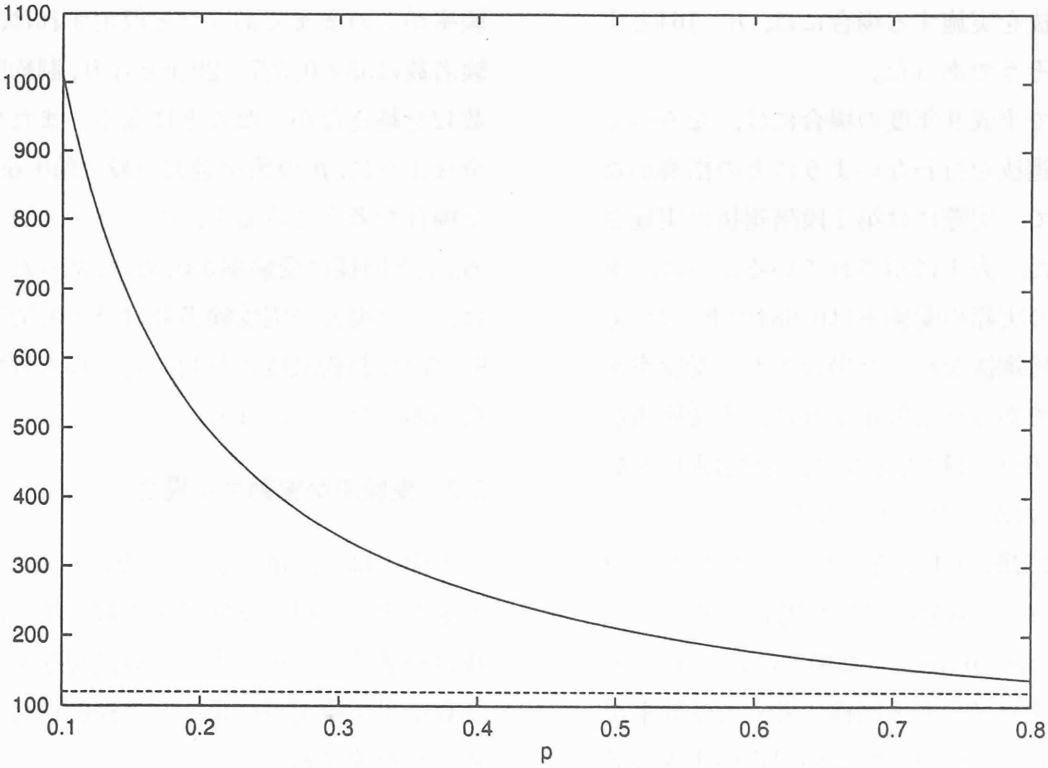

図 1: 受験率 p に対する第 1 段階選抜合格者数 n の変化例

る。これらの計算結果を見ると，2.3節で与えた方法によって第1段階選抜合格者数を決定しても，大きな問題はなさそうである。

［例4］ 平成9年度のA，B学科の場合

［例3］と同様な考え方によって，次に表1のA，B学科の平成9年度における第1段階選抜者数 n を決定し，実際のデータとの比較を行ってみよう。

上でも述べたように，平成9年度においては，大学入試センター試験の平均点に科目間格差があるとの理由により，可能な限り第1段階選抜を行わないようにとの指導があった。そこでA学科においてはこれを行わないことにしたが，B学科では2次試験で実験を行わせる必要があるため，多少の第1段階選抜を行わざるを得なかった。したがって下の表3に示すように，A，B学科の制約限度人数は，それぞれ $a_A=143$，$a_B=150$ と考えられた。この場合A学科では第1段階選抜を行わなかったので，受験予定者数，$\hat{c}_\alpha$ の値，それらの値の積として得られる予想受験上限，実際の受験者数およびそれと予想受験上限の差を，表3としてまとめた。ここで $\hat{c}_\alpha$ の値は，平成5年度から8年度までの過去4年分のデータを用いて，2.3節で与えた方法により計算した。

表 2. 平成5～7年度のデータから
平成8年度を推定した場合

	A 学 科	B 学 科
制約限度数 a	72	63
$\hat{c}_\alpha$ の値	0.601	0.797
推定人数 n	119	79
受験者推定値	59.5	42.7
差（余裕数）	+12.5	+20.3

表3　平成5～8年度のデータから
平成9年度を推定した場合

	A 学 科	B 学 科
制約限度数 a	143	150
$\hat{c}_\alpha$ の値	0.590	0.739
予想受験上限	84.4	110.9
実受験者数	89	77
差（余裕数）	−4.6	＋33.9

この表より，A学科においては僅かではあるが，かなり安全に見積もった予想受験上限を越えて，実際の受験者があったことが分かる。これは第1段階選抜を行わないと，それを行うことを想定して構築したモデルが当てはまらなくなる可能性のあることを示唆しているとも考えられ，解析に際して十分注意すべきであろう。

次にB学科では平成9年度においても第1段階選抜を行ったので，［例3］と同様な考察を行ってみよう。この場合 $a_B=150$，$\hat{c}_\alpha=0.739$であるから，推定人数は $n=a_B/\hat{c}_\alpha=150/0.739=203.0$となる。ところで平成9年度のB学科における受験率は0.513であったから，予想受験者数は203×0.513＝104.1と計算でき，余裕数は150－104.1＝45.9となる。したがってこの場合においても，2.3節で与えた方法は有効であると考えられる。なおB学科では，科目間格差のあった平成9年度においても第1段階選抜を実施したが，これは実験器具数の制約からやむをえず行ったものである。

［註］　科目の平均点にどの程度までの得点差があった場合には，第1段階選抜を行うべきでないという検討は，極めて難しい問題であろう。その1つの原因は科目間の得点差が本当に問題の難易度によるものなのか，その科目の受験層の質の違いによるものなのかが分離できない点にある。したがって，何らかの方法（例えばブートストラップ検定等）によって科目間の得点に統計的有意差が認められた場合には，第1段階選抜を極力避けた方がよいと考える。

4　おわりに

前節までにおいて，欠席者を考慮して第1段階選抜人数を決定する方法について，過去の受験率が比較的安定している場合と，かなり変動している場合に分けて考察を進めた。

そこでは，受験予定者が受験する確率 p は年度によらずすべての受験者に対して一定と仮定した。しかし試験の得点が高い場合は受験率が低くなる（または高くなる）といった，より現実的と考えられる仮定をおいた場合には，例えばロジスティック・モデルのような，さらに複雑なモデルを考える必要がある。しかし，表1のような限られたデータから多くの未知母数を含む複雑なモデルを推定することは，モデルの適合性や安定性などを検討する上で問題が多いと思われる。この点については，今後の検討課題としたい。

次に，過去の受験率がかなり変動していると考えられる場合については，受験率 p をベータ分布にしたがう確率変数と考え，分布のパラメータをモーメント法によりデータから推定する方法を提案した。この方法を表1のC学科やD学科のように，過去の受験率が比較的安定している場合に適用するとどのような結果が得られるであろうか。3.2節の［例3］および［例4］と同様な検討を行ったところ，C学科については $\hat{c}_a$ の値がうまく求められなかった。またD学科についても，平成8年度の予測はうまくいかなかった。現在その原因を検討中であるが，分散推定値 s^2を工夫しなければならない可能性がある。しかし平成9年度のD学科の解析においては，$\hat{c}_\alpha=0.870$，予想受験上限＝47.9，実受験者数＝29，余裕数＝＋18.9となり，妥当と思われる値が得られている。したがってこの点についても，

今後検討を進めなければならないと考えている。

最後に，2.3節で考察した問題は，定式化の仕方によってはベイズ統計におけるK群の確率の推定問題とも考えられる。今後Shigemasu & Mayekawa (1996)等の文献を参考にして，この問題についても検討していく予定である。

謝 辞

本研究に関して，大学入試センター研究開発部の前川眞一先生には大変有益な助言を頂きました。特に第4節の最後の点は，前川先生のご指摘によるものであります。ここに厚く御礼申し上げます。

参考文献

[1] 矢野一幸，大内俊二，田栗正章：大学入試における予備選抜倍率についての検討，行動計量学，17-2，25-33 (1990).

[2] K. Shigemasu and S. Mayekawa : A Bayesian hierarchical linear model with educational applications, J. Jap. Statist. Soc, 26-1, 1-23 (1996).