総合卒業試験による医師国家試験の 合否予測はどこまで可能か

平野光阳

入学する学生の質を高め、教育の質を向上させることは、6年間でストレートに医師となる者の数を 増やすとともに、将来に向けて有望な医師を誕生させることになるが、世間では医師国家試験(以下、 国試と呼ぶ。)の合格率を大学教育の1つの評価基準としているので,これを高めることにも無関心で はいられない。そこで、国試と密接な関係のある総合卒業試験(以下、総合卒試と呼ぶ。)の間の中か ら、項目反応の理論を応用して、何種類かの基準で、識別性能の高いものを選び、これらの間だけによ る得点で、合否の予測がどこまで可能か追究し、あまり満足すべき結果ではないが、総得点による場合 より予測力を高めることが出来た。

キーワード:総合卒業試験、医師国家試験、項目反応、識別性能、合否予測

1 はじめに

80年に1期生を迎えた本学は、今年10期生を送り出 し、これまでに千人近い医師を誕生させた。この間の国 試の大学別合格率による順位を見ると,92,93年は極め て上位であったが、昨年に続き今年も下位に低迷した。 世間ではこの合格率を大学教育の1つの評価基準として いるので、2年連続の不振を憂慮している向きも多い が、表1に示したように、6年間で卒業した者の比率が 開校以来一度も90%を下回ったことがなく、入学した 100人の内の何人が6年間でストレートに医師となった かを比較すれば、10年連続80人以上という記録を達成し た本学の教育は高く評価されるだろう。 1人でも多く国 試に合格させることだけが大学の教育目標でないことは 言うまでもないが、「入学する学生の質の向上に努め、

6年間のストレート卒業者を多くするとともに、医学部 を卒業することを前提に行われる国試の合格率を高める こと。」をいろいろある教育目標のうちの主要な1つと して捉えることには異論はあるまい。

国試の合格率を高める要因については、これまでに数 多くの研究成果を発表してきた。文献1)~4)及び 7) では、入学時の学力レベル (表1の2) 及び国試の 大学としての成績(表1のx)を年度の間で比較し、zと x の関係を追究した。また、文献 5) では、個人とし て見た場合の学内成績及び入試の成績が、国試の合否と どのようにかかわっているか追究した。しかし、入学時 の競争率が高くなると、統計学上の「選抜効果」6)の影 響が強く出るので、入試の成績と国試の合否の関係は見 られなくなるのが普通である。昨年は出身高校所在県、 年齢等の測定誤差のない入学時の属性や、選抜効果の影 響の少ない高校調査書の学習成績概評、出席状況及び課

表1	倍率·	入学問	寺の学力	りレ〜	ベル・	卒業者数	• 国試合	格率等の年度間の比較
		卒	-la Mi	卒	延		国	試

		20.	1111-) (-j-n	0-217	J/-	1 ///	3% = 17101	H - 1 ->	及同步起极	
7 224		中所		卒	卒業	卒 延		玉	絬		修正不合格者数
入学 年度	入学者数	実質 倍数	z —值	業年	者数	者 数	合格	不合格者数	合格率	全国合格率	$x = a + \frac{14.5 b}{11.11}$
T-12		111 35			7	(a)	者数	(b)	(d%)	(c %)	x = a + 100 - c
80	100*	3.6	1.46	86	90	10	84	6	93.3	86.6	16.5
81	100	2.3	1.50	87	97	3	94	3	96.9	86.2	6.2
82	100	3.1	1.70	88	96	4	88	8	91.7	81.2	10.2
83	100 *	2.6	1.47	89	91	9	80	. 11	87.9	88.0	22.3
84	100	2.9	1.61	90	97	3	86	11	88.7	82.9	12.3
85	100 *	1.0	1.28	91	93	7	85	8	91.4	84.3	14.4
86	100	3.8	1.47	92	90	10(8)	86	4	95.6	84.0	11.6**
87	100	9.6	1.66	93	92	8	90	2	97.8	90.1	10.9
88	100	6.0	1.63	94	93	7	80	13	86.0	86.2	20.7
89	100	7.1	1.64	95	93	7	82	11	88.2	86.0	18.4

*沖縄留学生1名を除く。** ()内の補正値8を使った。100-cは全国不合格率で,14.5は全国不合格率の88年ま での9年間の平均である。

山梨県中巨摩郡玉穂町山梨医科大学数学

(受付:1995年8月31日)

外活動等の評価によってグループ分けし、これらのグ ループ間で国試合格率に差があるか否か追究したとこ ろ, 多くの分類変数に関して期待を上回る結果が得られ た。^{8),9),10)} これらについてはさらに突っ込んで追究しているが、その成果は国立大学入学者選抜研究連絡協議会第16回大会で発表し、大学入試研究ジャーナル第6号に執筆中である。

入学する学生の質を高め、教育の質を向上させることは、単に国試の合格率を高めることのみならず、将教育の財で有望な医師を誕生させることにもつながり、教育の基本であろう。しかし、国試と言えども「一発勝負」の試験であるから、ストレート卒業者数を多くすれば、6年間でストレートに医師になる者の数が多くなる。そで、もし卒業試験の成績等によって国試の合本させにもう1年間勉強させることによって、6年間でストレートに医師になる者の数をあまり減らさずに、国試の合格率を飛躍的に高め、その順位を全国のトップに押し上げることも可能になる。特に、国試の模試とも言われている総合卒試によって、その予測がどこまで可能か、以下考察する。

2 卒業試験の成績と国試の合否の関係

国試の出題範囲になっている臨床科目等の卒業試験の成績と国試の合否の関係について、95年のデータを含めて考察する。国試の合否によって受験者(標本)を2群に分け、成績を順位で表し、小さい方の群(不合格者群)に属する n_1 人の順位数の和Rを求め、このRの値

が「両群の母集団分布が同一である。」という仮説の棄却域に落ちるか否かを調べる(順位和検定)ため、大きい方の群の人数 n_2 が十分大きいとき、Rの確率分布は、平均 μ_R 、標準偏差 σ_R の正規分布で近似できるので、

$$z_1 = \frac{R - \mu_R}{\sigma_R}$$

とする (標準化)。但し、

$$\mu_{R} = \frac{n_{1}(n_{1} + n_{2} + 1)}{2}$$

$$\sigma_{R} = \sqrt{\frac{n_{1}n_{2}(n_{1} + n_{2} + 1)}{12}}$$

また,両群の成績の平均値の差 x_1-x_2 が「母集団平均値の差 $\mu_1-\mu_2=0$ 」という仮説の棄却域に落ちるか否かを調べるため,標本標準偏差 s_1 , s_2 を母集団標準偏差 σ_1 , σ_2 の近似値として,

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{{S_1}^2}{n_1} + \frac{{S_2}^2}{n_2}}$$

を求め,

$$z_2 = rac{\overline{x}_1 - \overline{x}_2}{\sigma_{\overline{x}_1 - \overline{x}_2}}$$

とする。この場合も n_1 , n_2 が十分大きければ,母集団 分布が正規分布であるという仮定を除いても, z_2 の確率分布は近似的に標準正規分布に従う。いま,不合格者 群の標本数は十分大きいとは言えず,多少の誤差は避け られないが,標準正規分布では,関係式

$$P(z > 1.65) = 0.050$$

 $P(z > 1.96) = 0.025$

表 2 卒業試験(臨床科目等)の成績と国試の合否の関係

	r			<u> </u>	THE PROPERTY OF	「中田レベヤ	1 1 4 7		- BMV	百百の医	3 1015			
科目	9	0	9	1	9	2	9)3	9)4	9	5	平	均
	z_1	z_2	z_1	z_2	z_1	z_2	z_1	\mathcal{Z}_2	z_1	z_2	z_1	\mathcal{Z}_2	z_1	z_2
A	-0.40	-0.16	3.17	⑤ 3.76	1.95	1.83	2.42	2.06	1.40	1.51	2.45	2.47	1.84	1.86
В		22.78			0.53	0.32	2.29	2.22	1.32	1.37	2.55	2.69	2.26	2.35
C	2.22	32.76	43.45	43.99	0.50	0.58	2.29	4 2.69	2.66	3.20	0.41	0.44	1.89	2.11
D	0.02	0.18	⑤3.42	24.32	12.96	①3.62	0.83	0.84	43.79	43.48	2.06	2.37	2.26	⑤ 2.51
E	1.05	1.17	1.60	1.54	1.62	1.56	1.59	1.78	⑤3.58	43.48	3.11	⑤3.13	2.29	2.40
F	-1.30	-1.44	2.94	2.92	1.97	1.88	0.82	0.78	2.40	2.38	43.48	33.55	1.92	1.90
G	1.43	1.10	0.07	-0.15	-0.45	-0.69	1.49	1.98	1.10	1.09	-0.07	0.12	1.02	1.26
Η	1.43	1.32	1.34	1.58	1.90	1.62	1.14	1.35	2.98	3.25	1.21	1.30	1.86	1.96
I	1.29	1.44	1.69	1.61	32.63	2.62	②2.68	32.79	15.14	15.26	②3.73	23.84	33.06	33.22
J	32.66	⑤ 2.61	0.92	0.90	2.34	42.33	1.16	1.33	2.05	2.33	1.90	1.85	2.22	2.31
K	1.96	0.84	1.77	1.71	0.51	0.56	0.23	0.43	0.19	0.56	0.09	0.47	1.14	1.14
L	2.31	2.43	2.17	2.15	4 2.47	⑤ 2.20	1.33	1.02	2.63	2.86	⑤3.19	2.97	⑤2.44	2.40
\mathbf{M}	1.77	1.90	1.56	1.57	1.83	1.74	1.03	0.91	3.38	3.29	2.36	2.21	2.36	2.26
N	1.40	1.41	0.61	0.72	-1.91	-2.48	0.73	0.75	1.93	2.02	-0.21	-0.16	0.59	0.64
О	2.01	2.31	3.07	3.51	-0.11	-0.05	⑤ 2.50	①3.50	3.24	3.42	2.44	2.53	1.99	2.20
P	⑤ 2.55	2.36	1.47	1.42	0.51	0.71	0.65	0.70	3.24	3.32	3.04	2.74	1.82	1.71
Q	-0.94	-0.81	0.56	0.46	⑤ 2.41	1.99	0.20	0.50	1.82	1.78	0.89	1.04	0.91	0.75
R	1.79	1.67	23.67	3.74	1.52	1.47	4 2.52	2.54	2.61	3.17	14.04	14.22	42.65	42.78
S	1.89	1.39	2.81	2.97	0.42	0.22	0.91	0.85		0.87	0.77	0.90	1.54	1.59
T										34.01		43.40	①3.44	①3.69
U	22.67	4 2.68	33.47	34.29	2.21	1.92	22.68	⑤2.68	24.46	24.19	3.16	3.05	23.24	23.32

 $A \sim S$ は臨床科目等19科目と乱数表を用いて対応付けたもので、T は総合卒試、U は全科目の平均である。横の平均は86年~95年の平均であるが、87年及び88年を含まない科目が一部ある。①~⑤は値の大きい方からの順位である。

	1	2	3	4	5
\mathbf{E}	F F. F F F	F	F. F	F F.	F
F	F F F F F F	F	F	F F F	
	F F F F F F. F				
	F F. F F. F.				
	F F F F. F F F				
	F. F F. F				
U	$F. F. \ldots F. \ldots FF.$. F F	F	F	F F
	図1 国	試不合格者 (F)	の下から数えた順位	(その1)	

P(z>2.33)=0.010 P(z>2.58)=0.005P(z>3.08)=0.001

が成り立つ。但し、zは z_1 , z_2 の総称である。

 z_1 , z_2 は86年卒から求めてあるが,90年以降の6年分を表2に掲載した。なお,対象は留年経験者も含めてその年に卒業した者全員である。10年間を通して考えれば不合格者群も十分な標本数になるが,異なる試験の成績を合算するわけにはいかないので,10年間の平均を示した。多少ラフな検定になるが,年ごとの変動があまり大きくなければ,標本数が10倍になったと考えられるので,平均のzに代えて $\sqrt{10}z$ を検定に用いるとよい。

T(総合卒試)についての Z(以下,単にTと呼 ぶ。)が平均1位で、毎年4位以内にあるのは合点のい くところであるが、Tを含む全科目の平均UがTを下回 り、最近では5位以内に入っていない年もあるのを意外 に思う向きは多かろう。原因の1つとして, 前年に卒延 となった者がこの1年間に実力を付けても、昨年際どく 合格した科目(Tは含まれない。)の点数は変らないと いうことが考えられるが、Uの極端に悪い者の多くが卒 延になり、この比較の対象外(合否どちらの群にも入ら ない。)になっていることを見逃すわけにはいかない。 すなわち、「選抜効果」が働いているのである。ちなみ に、6年次での卒延者が皆無で、多くの不合格者の出た 昨年はUがTを上回っている。また、科目別に見ると、 平均3位の I は昨年1位で、5.0 を超える驚異的な値と なっており、今年も2位である。また、94、95年には科 目 D, E, L, M, R などが大きく, 半数以上の科目で有 意水準 0.025 (片側) の棄却域に入っている。長期的に 見ると,ある時期に大きく変った科目や年ごとに変動の 激しい科目もあるが、多くの科目でそれぞれの傾向が見 られる。

不合格者11人の下からの順位を図1に示した。不合格者のTの平均順位(以下,平均順位を省略)は76.1(95人中),Uは72.7,Rは79.6,Iは77.2で,下から10番までに入っている不合格者の数は T, U がともに 3, R が 6, I が 5 である。従って,T 又は U で下から10人を卒延にすると,受ければ合格した者 7 人を落とすことになり,この中には留年経験者もいるが,合格率が多少上がる代りに, 6 年間でストレートに医師になる者の数が大幅に減少する。R や I の Z は,T や U と違い,年による変動が大きいから,R や I の成績でこれを行うわけにはいかない。なお,昨年の不合格者数は17と大きい

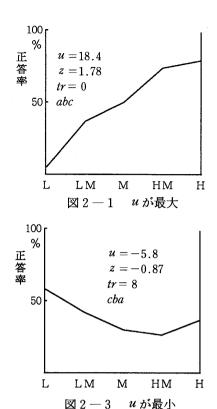
ので、下から10番までに、Tで6人、Uで7人、Iで8人が入っている。しかし、不合格者数が今年と同じ11でも、Tの下から7番までが全員不合格という年もある。

3 総合卒業試験の各問の実力識別性能

項目反応¹¹⁾の理論を応用して、総合卒試の各間(項目)が学生の実力差をどのように識別しているか、その実力識別性能を測るため、卒延となった 1 人を含め、96 人の受験者を次の 3 通りの方法でグループ分けする。第 1 のグループ分けは最下位(L, 19人)、下位(LM、19人)、中位(M, 20人)、上位(HM, 19人)、最上位(H, 19人)の 5 群、第 2 は下位(C, 31人)、中位(B, 33人)、上位(A, 32人)の 3 群、そして第 3 は国試不合格者群(卒延の 1 人を含む。)と合格者群の 2 群である。次に、項目ごとに各群の正答率を求め、この正答率を各グループ分けの下の群から上の群に向って、 x_1, x_2, \dots, x_n とする。11)の論文で高野氏は、「実際 $0 \le x_1 \le x_2 \le \dots \le x_n \le 1$

を満たさない項目もあるが、そのような項目は"異常"なものとして、分析の対象から除く。」としているが、センター試験の場合と違い、第1の5群の場合、321間(正答率100%の3間を除く。)中この不等式をすべて満たすものは12%にも満たない38間である。従って、これを満たさない間を対象から除くわけにはいかない。センター試験とこのように大きな相違が見られる原因として、センター試験は各群の人数が何万人であるのに対し本試験は20人足らずであること、センター試験は項目数が50前後であるのに対し本試験は321もあるので、1つの項目の総点への寄与率が小さいことが考えられる。なお、後者の事実は分析にとって好ましいことである。

続いて、実力識別性能を測る尺度を 4 種類導入する。第 1 の尺度は、L,LM,M,LH,Hに対する x_1 , x_2 , x_3 , x_4 , x_5 の間で「i < j かつ $x_i > x_j$ 」が成り立っていれば転位が起きているとし、すべての i, j の組合せに対する転位の数(転位数と呼ぶ。)である。転位数ごとの項目数は表 6 にあるが、1 及び 2 の場合が最も多く、それぞれ76項目(23.7%)である。言うまでもなく、転位数の少ない項目程識別性能が高いと見なす。第 2 の尺度は C,B,A に対する $x_1 = c$, $x_2 = b$, $x_3 = a$ の大きさの順序で、abc から cba まで 6 通りあるが、abc に近い程識別性能が高く、cba に近い程低いとみなす。第 3 の尺度は、単に大小関係だけではなくその差も考慮

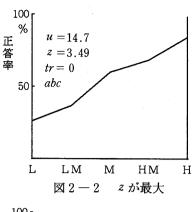


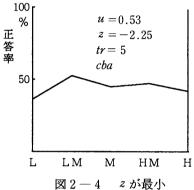
した x_i (%) のi に対する回帰直線の傾きu (正答を 1, 誤答を 0 として,総点との間の相関係数を用いても同様) である。もちろん,u が大きい項目程識別性能が高いと見なす。第4 の尺度は国試の合格者群と不合格者群の正答率の差を標準化したもので,

$$z = \frac{p_1 - p_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

で表される。但し、クは96人全員の正答率である。

zが大きい項目程,不合格者を予測する性能が高い項目であるということになるが,言うまでもなくこれを予測に用いることは出来ないので,他の3つの各尺度とzの関係を調べ,その関係が強ければ,その尺度による識





別性能の高い項目を予測に用いることにする。321の問の中から代表的な4つを選び,各群と正答率の関係をグラフで見ると,図2-1 (uが最大),図2-2 (zが最大)の場合は,tr(転位数)=0で,いずれも識別性能が高いが,図2-3及び図2-4の項目は全く識別力がないか負の識別力で,妥当な問であったかどうか疑わしい。

4 識別性能を測る尺度の相互の関係

表 3 及び表 4 は回帰直線の傾き u 及び標準化された正答率の差 z と正答率 t (%) の関係を示したものである。また,図 3 は t と u, t と z 及び z と u の関係をグラフに表したものである。正答率が10%以下の問は 5 問

_					
主り	同帰直線の傾き	(11)	人工女女	(+ 0/ \	か関する
		(u)		(1 70)	U J DELITE

	$t \leq 1$	$10.0 < t \le 2$	$20.0 < t \le$	$30.0 < t \le$	$40.0 < t \le$	$50.0 < t \le$	60.0< t ≤	$70.0 < t \le$	$80.0 < t \le$	90.0 < t	sum
$u \leq -5.0$	0	0	0	1	Ó	1	0	0	. 0	0	2
$-5.0 < u \leq -2.5$	0	0	2	2	1	1	1	1	0	0	8
$-2.5 < u \le 0.0$	3	5	7	4	4	2	3	4	2	8	42
$0.0 < u \le 2.5$	2	5	4	4	4	2	5	6	7	18	57
$2.5 < u \le 5.0$	0	5	2	9	6	4	7	11	10	9	63
$5.0 < u \le 7.5$	0	1	1	6	8	8	8	6	10	2	50
$7.5 < u \le 10.0$	0	2	1	3	13	10	14	11	5	0	59
$10.0 < u \le 12.5$	0	0	1	1	2	6	6	1	4	0	21
$12.5 < u \le 15.0$	0	0	1	2	0	4	4	3	0	0	14
15.0 < u	0	0	0	1	1	2	3	1	0	0	8
sum	5	18	19	33	39	40	51	44	38	37	324
mean u	-0.42	2.22	1.83	4.11	5.99	7.53	7.45	5.65	5.21	1.69	5.04
s. d. <i>u</i>	1.14	3.10	5.02	5.25	4.45	5.09	4.51	4.41	3.26	1.86	4.71

mean t = 59.7, s.d. t = 23.8, r(相関係数) = 0.066

	<i>t</i> ≤	$10.0 < t \le 2$	$20.0 < t \le 3$	$30.0 < t \le 3$	$40.0 < t \le$	$50.0 < t \le$	$60.0 < t \le $	$70.0 < t \le 3$	$80.0 < t \le$	90.0< t	sum
$z \leq -1.5$	1	1	2	4	2	0	3	0	0	0	13
$-1.5 < z \le -1.0$	0	1	2	1	2	0	3	0	1	2	12
$-1.0 < z \le -0.5$	2	2	1	2	3	7	0	7	2	5	31
$-0.5 < z \le 0.0$	0	2	2	6	4	8	2	4	6	11	45
$0.0 < z \le 0.5$	2	3	2	4	5	4	8	6	4	2	40
$0.5 < z \le 1.0$	0	4	5	4	8	2	12	7	9	4	55
$1.0 < z \le 1.5$	0	2	4	6	6	5	6	7	4	2	42
$1.5 < z \le 2.0$	0	3	1	2	6	9	5	6	4	3	39
$2.0 < z \le 2.5$	0	0	0	3	3	1	6	2	5	1	21
2.5 < z	0	0	0	1	0	4	6	5	3	7	26
sum	5	18	19	33	39	40	51	44	38	37	324
mean z	-0.71	0.39	0.22	0.41	0.60	0.76	0.96	0.86	0.95	0.61	0.68
s. d. <i>z</i>	0.89	1.00	1.10	1.26	1.11	1.19	1.19	1.12	1.08	1.29	1.18

表4 標準化された正答率の差(z)と正答率(t%)の関係

r(相関係数)=0.168

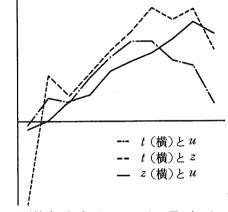


図3 正答率 (t) とu, t とz 及びz とu の関係

だけであるが、それらの平均はu, zともに負になっており、いずれも識別性能はない。t の上昇に伴ってuもz も増加し、 $50.0 < t \le 70.0$ を満たす問の過半数で7.5 < u となっているが、当然のことながら、t がさらに大きくなるとu は減少の傾向を示し、90%以上では37 間中35間で $-2.5 < u \le 5.0$ となっている。これに対し、z の方は $60.0 < t \le 90.0$ で最も大きく、90.0 < t となってもあまり小さくはならない。卒延者1 人を不合格者と見なした国試合格率が87.5%であることを考えると、これも納得出来よう。なお、u = f(t) のグラフは上に凸な放物線に近いが、u もz も正答率別の標準偏差が比較的大きい。

表 5-1 回帰直線の傾き(u)と標準化された正答率の差(z)の関係

及 1 国											
	z ≤	$-1.5 < z \le -1.5$	$-1.0 < z \le -$	$-0.5 < z \le$	$0.0 < z \le$	$0.5 < z \le$	$1.0 < z \le$	$1.5 < z \le$	$2.0 < z \le$	2.5 < z	sum
$u \leq -5.0$	0	0	2	0	0	0	0	0	0	0	2
$-5.0 < u \le -2.5$	3	2	0	2	0	1	0	0	0	0	8
$-2.5 < u \le 0.0$	6	3	9	10	4	4	5	1	0	0	42
$0.0 < u \le 2.5$	3	5	7	13	6	9	4	5	1	4	57
$2.5 < u \le 5.0$	1	1	7	11	13	11	9	4	1	5	63
$5.0 < u \le 7.5$	0	1	2	6	9	10	7	8	5	2	50
$7.5 < u \le 10.0$	0	0	4.	3	4	15	10	10	7	6	59
$10.0 < u \le 12.5$	0	0	0	0	4	2	2	7	3	3	21
$12.5 < u \le 15.0$	0	0	0	0	0	3	3	3	2	3	14
15.0 < u	0	0	0	0	0	0	2	1 ,	2	3	8
sum	13	12	31	45	40	55	42	39	21	26	324
mean u	-0.77	0.09	1.88	2.54	4.68	5.56	6.38	7.67	9.25	8.24	5.04
s. d. <i>u</i>	1.93	3.05	3.74	2.95	3.65	4.06	4.60	4.26	3.75	5.08	4.71

r(相関係数)=0.568

表 5 - 2

	$z \leq 0$	sum		
$u \leq 2.5$	62	24	20	106
	(32.4)	(31.4)	(42.3)	
$2.5 < u \le 7.5$	29	43	41	113
	(34.5)	(33.4)	(45.1)	
7.5 < u	7	28	67	102
:	(31.1)	(30.2)	(40.7)	
sum	98	95	128	321

 $\chi^2 = 80.5$, r(クレーマーの関連係数) = 0.354

zとuの関係を表すグラフは直線に近く、 $u \le 0.0$ を満たす52問の71%に当る37問で $z \le 0.0$ 、10.0 < u を満たす43問はすべて 0.0 < z で、その79%に当る34間で 1.0 < z で ある。また、相関係数 r = 0.568 (n = 324)、 $\chi^2 = 80.5$ ($\nu = 4$)、 $\rho \nu - \nu - \nu$ の関連係数は 0.354 である。なお、 $\nu = 4$ で独立性の検定をするとき、確率に関する関係式

 $P(\chi^2 > 9.49) = 0.050$ $P(\chi^2 > 11.14) = 0.025$ $P(\chi^2 > 13.28) = 0.010$ $P(\chi^2 > 14.86) = 0.005$ $P(\chi^2 > 18.47) = 0.001$

が成り立つ。

以上から、合格率が90%前後の国試の模試としては正答率が60%~90%の問が望ましいこと、 *u* の大きな問に関連した事項は特に注意して勉強しておく必要があることなどが分かる。

表 6 は転位数と z の関係,表 7 は a, b, c の順序と z の関係を示したものである。転位数が 0 又は 1 の問の 82%に当る94問で 0.0 < z であるのに対し, 6 以上の 間では22問中17問(77%)で $z \le 0.0$ で, z は転位数 の増加に伴って減少している。 v=4 で $x^2=54.1$, r

表6-1 転位数と標準化された正答率の差(z)の関係

									, ,-	_ `	
	z	1	2	3	4	5	6	7	8	9	sum
	0	0	5	1	1	7	6	7	4	7	38
	1	2	2	10	9	15	8	11	10	9	76
	2	2	4	7	14	7	14	15	5	8	76
転	3	0	5	10	8	16	4	3	1	2	49
位	4	4	5	9	3	6	7	1	1	0	36
数	5	8	4	3	3	4	1	1	0	0	24
	6	6	2	2	1	0	2	1	0	0	14
	7	2	3	0	1	0	0	0	0	0	6
	8	1	1	0	0	0	0	0	0	0	2
su	m	25	31	42	40	55	.42	39	21	26	321

1. $z \le -1.0$, 2. $-1.0 < z \le -0.5$,, 9. 2.5 < z

表 6 - 2

		$z \leq 0$.	sum		
	0, 1	20	32	62	114
転		(34.8)	(33.7)	(45.5)	
位	2, 3	28	45	52	125
数		(38.2)	(37.0)	(49.8)	
双	4 ~	50	18	14	82
		(25.0)	(24.3)	(32.7)	
	sum	98	95	128	321

 $\chi^2 = 54.1$, r(クレーマーの関連係数) = 0.290

表 7-1 上位群 (a)・中位群 (b)・下位群 (c) の 正答率の順序と標準化された正答率の差(z) の関係

z	1	2	3	4	5	6	7	8	9	sum
abc	4	9	15	18	30	24	30	19	17	166
acb	5	7	12	12	9	5	2	0	2	54
bac	0	6	7	5	11	12	6	1	7	55
bca	2	3	3	1	1	1	0	1	0	12
cab	12	- 5	3	2	3	0	1	0	0	26
cba	2	1	2	2	1	0	0	0	0	8
sum	25	31	42	40	55	42	39	21	26	321

1. $z \le -1.0$, 2. $-1.0 < z \le -0.5$,, 9. 2.5 < z

表 7 - 2

	$z \leq 0$	$0 < z \le 1$.	0 < z	sum
abc	28	48	90	166
	(50.7)	(49.1)	(66.2)	
acb, bac	37	37	35	109
	(33.3)	(32.3)	(43.5)	
bca, cab	33	10	3	46
cba	(14.0)	(13.6)	(18.3)	
sum	98	95	128	321

 $\chi^2 = 60.9$, $r(\rho \nu - \neg \neg \sigma$ 関連係数)=0.308

(クレーマーの関連係数,以下同様)=0.290 であるから,両者が独立でないことは確かである。また,3群に分けた場合は過半数に当る166間で正答率が abc の順になっているが,高野氏による "異常" な問も48%ある。転位数が同じでも,acb ($z \le 0.0$ の問が44%) の方がbac (同 24%) より,cab (同 77%) の方がbca (同 67%) よりz の平均が小さい。すなわち,c が上に来るとz が負になりやすいという傾向が見られる。なお, x^2 は前の場合とあまり違わない。

表 8 は転位数と abc の順序,表 9 は転位数と u,表 10は abc の順序と u の間の関係を示したものである。 abc で転位数が 6 の問,cba で転位数が 4 の問も無いわけではないが,概して abc では転位数が $0 \sim 2$, acb と bac では $2 \sim 4$, bca と cab では $4 \sim 6$, cba では $6 \sim 7$ に多く,全体的には妥当な分布であるが,c が上に来た方が転位数も多くなる傾向がある。 u との関係は,予想された通り,z との関係よりはるかに顕著で,いず

表 8-1 転位数 2a, b, c の順序の関係

			1,504 1		, , ,	- /104/3	- DUDI	•
		abc	acb	bac	bca	cab	cba	sum
	0	35	1	2	0	0	0	38
	1	61	8	7	0	0	0	76
	2	41	15	19	0	1	0	76
丰二	3	20	13	15	1	0	0	49
転位数	4	5	12	8	4	6	1	36
数	5	2	4	4	3	10	1	24
	6	2	1	0	2	6	3	14
	7	0	0	0	1	2	3	6
	8	0	0	0	1	1	0	2
S	um	166	54	55	12	26	8	321

表 8 - 2

	abc	acb , bac	その他	sum
0, 1	96	18	0	114
	(59.0)	(38.7)	(16.3)	
2, 3	61	62	2	125
	(64.6)	(42.4)	(17.9)	
4 ~	9	29	44	82
	(42.4)	(27.8)	(11.8)	
sum	166	109	46	321

 $\chi^2 = 188.9$, $\gamma(\rho \nu - \tau - \sigma)$ 関連係数)=0.542

表 9-1 転位数と回帰直線の傾き(u)の関係

	и	1	2	3	4	5	6	7	8	9	sum
	0	0	0	5	2	5	11	4	7	4	38
	1	0	0	5	13	17	20	12	6	3	76
4	2	0	1	11	19	20	19	4	1	1	76
転	3	0	1	12	19	7	9	1	0	0	49
位	4	0	9	18	8	1	0	0	0	0	36
	5	2	15	5	2	0	0	0	0	0	24
数	6	3	10	1	0	0	0	0	0	0	14
	7	3	3	0	0	0	0	0	0	0	6
	8	2	0	0	0	0	0	0	0	0	2
S	um	10	39	57	63	50	59	21	14	8	321

1. $u \le -2.5$, 2. $-2.5 < u \le 0.0$,, 9. 15.0 < u

表 9 - 2

		30 0							
	$u \le 2.5 < u \le 7.5 < u$								
0,1	10	37	67	114					
転	(37.6)	(40.1)	(36.2)						
位 2,3	25	65	35	125					
	(41.3)	(44.0)	(39.7)						
数 4~	71	11	0	82					
	(27.1)	(28.9)	(26.1)						
sum	106	113	102	321					

 $\chi^2 = 172.0$, $r(\rho \nu - \tau - \sigma)$ 関連係数)=0.518

表10-1 abc の順序と回帰直線の傾き(u)の関係

и	1	2	3	4	5	6	7	8	9	sum
abc	0	3	15	26	32	51	19	12	8	166
acb	- 0	5	21	16	8	3	1	0	0	54
bac	0	4	12	21	10	5	1	2	0	55
bca	3	6	3	0	0	0	0	0	0	12
cab	4	16	6	0	0	0	0	0	0	26
cba	3	5	0	0	0	0	0	0	0	8
sum	10	39	57	63	50	59	21	14	8	321

1. $u \le -2.5$, 2. $-2.5 < u \le 0.0$,, 9. 15.0 < u

表10-2

	$u \leq 2$	sum		
abc	18	58	90	166
	(54.8)	(58.4)	(52.7)	
acb, bac	42	55	12	109
	(36.0)	(38.4)	(34.6)	
bca, cab	46	0	0	46
cba	(15.2)	(16.2)	(14.6)	
sum	106	113	102	321

 $\chi^2 = 167.3$, r(クレーマーの関連係数) = 0.511

れの 2 者間でも $167 < x^2 < 189$, 0.5 < r を満たしている。 $0 < x^2 < 189$, 0.354 の $x < x^2 < 189$ の間の相関係数が 0.354 の $x < x^2 < 189$ の複雑な $x < x^2 < 189$ の複雑な $x < x^2 < 189$ の複雑な $x < x^2 < 189$ の $x < x^2 < x^2 < 189$ の $x < x^2 < 189$ の $x < x^2 <$

表11 正答率 (t%) に制限を加えた場合の x^2 及びクレーマーの関連係数 (r) の比較

	0.0 < t	<100.0	5.0 <t< th=""><th><95.0</th><th colspan="3">10.0 < t < 90.0</th></t<>	<95.0	10.0 < t < 90.0		
	X 2	r	X 2	r	X 2	r	
$u \geq z$	80.5	0.354	78.5	0.357	68.0	0.347	
転位数とを	54.1	0.290	55.9	0.301	44.0	0.279	
abc の順序と z	60.9	0.308	60.4	0.313	56.2	0.316	
転位数とabcの順序	188.9	0.542	183.9	0.546	163.7	0.539	
転位数とи	172.0	0.518	197.6	0.566	189.9	0.580	
abc の順序と u	167.3	0.511	179.7	0.540	179.3	0.564	

の代りに、転位数あるいは abc の順序を用いても、結果にそう大きな違いはなかろう。

表 5-2 以降は正答率100%の問を除く321問を対象にしたものであるが、95%以上及び5%以下を除く308問、90%以上及び10%以下を除く282問を対象にしても、表11に示したように、それ程大きな違いは見られない。

表12 各条件によって選ばれた間による点数と国試の合否の関係

	The property of the property o										
条件	X 2	z_1	z_2	a	b	c	d	mfR	means	meanf	配点
total	12.18	3.58	3.42	47	37	0	11	76.0	197.1	171.1	324
$2.5 \le u$	9.46	3.73	3.77	49	35	1	10	77.2	133.9	105.7	215
5.0 < u	8.12	3.75	3.85	46	38	1	10	77.3	94.4	69.5	152
$7.5 \le u$	8.12	3.82	3.87	46	38	1	10	77.9	63.1	43.2	102
転位数3	12.18	3.82	3.86	47	37	0	11	77.9	153.5	124.4	239
転位数 2	12.18	3.95	3.97	47	37	0	11	78.9	126.9	101.0	190
転位数 1	8.99	3.50	3.55	48	36	1	10	75.4	79.0	63.0	114
acb, bac	12.18	3.77	3.76	47	37	0	11	77.5	174.2	144.8	275
abc	8.54	3.73	3.83	47	37	1	10	77.2	107.4	84.6	166
0.5 < z	12.71	4.86	5.62	48	36	0	11	86.0	116.8	82.5	183
1.0 <z< td=""><td>12.18</td><td>5.06</td><td>6.06</td><td>47</td><td>37</td><td>0</td><td>11</td><td>87.6</td><td>83.3</td><td>54.2</td><td>128</td></z<>	12.18	5.06	6.06	47	37	0	11	87.6	83.3	54.2	128

転位数 1 は $0 \sim 1$, 2 は $0 \sim 2$, 3 は $0 \sim 3$, acb , bac は abc も含む。 a:合格で上位者,b:合格で下位者,c:不合格で上位者,d:不合格で下位者,mfR:不合格者の順位の平均,means:合格者の平均点,meanf:不合格者の平均点

	1	2	3	4	5
total	F. F F F	. F. F F	. F F F	· F	
2.5 < u	$. F F \ldots F \ldots F.$	F. FF F. F.	$\dots F \dots .$. F
	F F. F. F F F.				
	F F F F F F				
	F. F. F F F				
転位数 2	F F F F F.	. F FF F.	F F	7	
	$F. F. F. F. \dots FF. \dots$				
	. F F F F F				
	FFF F. F. F				
	FFFFFFFF.F				
1.0 < z	FFFFFFFFFF		$\dots F \dots .$		
	図 4 国	試不合格者(F)の ⁻	下から数えた順	位(その2)	

5 識別性能の高い問による国試の合否予測

識別性能の高い問を(1) 2.5 < u, (2) 5.0 < u, (3) 7.5 < u, (4) 転位数 3 以内, (5) 転位数 2 以内, (6) 転位数 1 以内, (7) abc, acb, bac, (8) abc, (9) 0.5 < z, (10) 1.0 < z の10種類の条件をそれぞれ満たすものとして選び、卒延者 1 人を除く95人について、各条件を満たす問による10種類の点数及びその点数による順位を求めた。(1) の条件を満たす間は全体の66.4%, (2) の間は46.9%, (3) の間は31.5%, (4), (5), (6) の間はそれぞれ73.8, 58.6, 35.2%, (7) の間は84.9%, (8) の間は51.2%, また (9) 及び (10) の間は56.5%及び39.5%である。

これらの点数について, 卒業成績による場合と同様 に、 χ^2 ($\nu=1$), 順位和及び平均値の差の検定に必要 な値を表12に示した。 χ^2 は独立であるという仮説を棄 却するのにいずれも十分大きな値となっているが、不合 格者で上位群に属する者が1人いるか否かですべてが決 まるので,(1)~(10)の比較には不向きである。いま, z₁ 及び z₂ に注目すると, (6) の場合の z₁を除いて, すべて の問を使った場合より大きくなっている。ルに関する条 件で問を制限した場合は、制限を厳しくする程とは大き くなっている。転位数による場合は(5)の場合が最高 で、(3) の場合も上回っている。(7) の場合と(8) の場合で はほとんど差が見られない。(9)及び(10)の場合は、当然 のことながら、 z が極めて大きくなっている。なお、 total の z が表 2 の T のものとわずかに異なるのは、前 者は素点を用い、後者は4捨5入によって100点満点に 換算したものを用いて計算しているからで、図1と図4 での違いも同じ理由による。

図4を見ると、uに関する条件で問を制限した場合は、不合格者の1人が大きく順位を上げてしまうため、平均順位がほとんど下がらないが、下から10番までに入る者の数は $3\sim5$ で、特に(3)の場合は、下から6番までに不合格者4人が入る。転位数では、(5)の場合に大幅に順位を上げる者がいないため、平均順位は最も下がるが、下から10番までには4人である。(8)の場合は下から3番までがすべて不合格者となるが、全体としてそ

の順位はほとんど下がらない。参考までに、(9)及び(10) の場合は、前者では下から7番まで、後者では下から8番までが見事に全員不合格者となる。従って、どの問がこの条件を満たしているかの情報は、次年度に受験する者にとって役立つであろう。

このように、いずれの順位も国試の合否と一致するものではない上に、その年に何人の不合格者が出るかの予測方法も確立せねばならないので、不合格者を高い確率で予測することは難しいが、引き続き改良に努める所存である。しかしながら、在学中の成績の極めて良くない者が国試に合格することは、まぐれ合格と考えられ、優れた医師を育成するという立場に立てば、7.5< u を満たす問による点数で下から6人程度を卒延にさせるのも一案であろう。

なお、総合卒試の成績と国試の合否が一致しないのは、前者の問だけに問題があるとは一概に言えないだろう。国試の間についても、既に実施されているのかも知れないが、項目反応の理論等を適用して、それが適切なものであったか否か調べてみることが望まれる。しかし、このいずれもが適切な問ばかりであったとしても、両者の結果が完全に一致することはなかろう。試験の結果は受験者の精神面や体調など試験時のいろいろなコンディションに左右されるからである。

謝 辞

本学在職中は入試の追跡調査・研究における共同研究 者であり、本論文の原稿に目を通されて貴重なご意見を 下さった、恵泉女学園学園長の川田殖先生、データの整 理、コンピュータへの入力、ワープロによる原稿作成の 一切を担当して下さった、入学者選抜方法研究委員会研 究補助員の三澤恵さんに、日ごろのご支援と合せて、感 謝の意を表したい。

文 献

1) 平野光昭: (1992) 面接の評価・学内成績・医師国 家試験の合否の関連。大学入試研究ジャーナル,第 2号,58~64

- 2) 平野光昭: (1992) 入学時の平均的学力及び専門教育と医師国家試験の合格率の関連。山梨医科大学紀要,第9巻,84~92
- 3) 平野光昭: (1993) 医師国家試験の合格率を高める 要因一受験機会の複数化・入学時の学力レベル・大 学教育一。大学入試研究ジャーナル,第3号, 23~30
- 4) 平野光昭: (1993) 国立大学の受験機会と入学者の 学力レベル及び同レベルと医師国家試験の合格率の 関係。大学入学者の特性と選抜方法との関連につい ての追跡調査研究(平成4年度科学研究費補助金に よる研究),研究成果報告書,149~156
- 5) 平野光昭: (1993) 卒業試験の成績及び入試成績等 と医師国家試験の合否の関係―主成分分析―。山梨 医科大学紀要,第10巻,69~78
- 6) 平野光昭: (1993) 国立大学の入試に関する常識と 非常識。名古屋大学教育学部紀要―教育心理学科

- 一, 第40巻, 4~14
- 7) 平野光昭: (1994) 医師国家試験の大学としての成績を高める入試及び他の要因一主成分分析—。大学入試研究ジャーナル,第4号,6~13
- 8) 平野光昭: (1994) 医師国家試験の合否と入学時の 属性及び高校調査書の内容の関係―どのような学生 を入学させれば国試の合格率が高まるか―。山梨医 科大学紀要,第11巻,29~38
- 9) 平野光昭: (1995) 入試成績・入学時の属性・学内 成績と医師国家試験の合否の関係。大学入試研究 ジャーナル, 第5号, 39~49
- 10) 平野光昭: (1995) 追跡調査の理論と実際―追跡調査でこんなにいろいろなことが分かる―。大学入試研究の動向,第12号,印刷中
- 11) 高野文彦: (1992) 試験の評価方法としての項目反 応の応用。大学入試研究ジャーナル,第2号,1 ~13

Abstract

How far is it possible for any student to predict the Pass-rate of National Examination for Medical License (NE) with the Aid of the Comprehensive Graduation Test (CGT)?

Teruaki HIRANO

The elevation of the quality of successful applicants and the improvement of the contents of education result in the increase of the number of medical students who are to go straight on to doctors in due course of six years and the birth of promising doctors. However, as the pass-rate of NE is often considered as a landmark of the contents of medical education in the universities and colleges, we cannot disregard the elevation of the Pass-rate of NE. Therefore we, applying the theory of the item response and using several kinds of standard selected from CGT particular questions which seem to serve as the decisive factors to distinguish between the able student and the unable one. Then we tried to find with the aid of the particular scores in these questions how far it is possible for any student to predict the pass-rate of NE. The result was not so satisfactory, but it was possible to improve the power of prediction more effectively than doing it with the aid of the whole scores in the whole CGT.

Department of Mathematics