

## 論述式テストの研究(2) : 小論文採点の集計法

阿久津 洋 巳\*

(2017年2月15日受理)

Hiromi AKUTSU

### A Study of Essay Tests (2) : Processing for Raw Scores

本論文は入学試験に使われる小論文の採点時の集計方法を検討した。小論文テストでは、採点者間の得点(素点)の分散が評価の公平性に影響する。この問題の解決方法は素点を $z$ 変換して分散を調整することにあるが、素点の平均と分散調整済み得点の平均の違いを明らかにするために数値シミュレーションを使用した。さらに現実場面に近い条件で、素点の分散の違いが順位に及ぼす影響を推測した。これらの例示を通して、素点を $z$ 変換する方法がもつ利点を明確にした。

#### はじめに

小論文は作文能力や「学力」、「能力」あるいは人物を推測する道具として、入学試験を初めとして様々な試験において長い間使われてきた。一方では、小論文が適切に期待される機能を果たすか(言いかえれば測定用具としての信頼性と妥当性があるのか)という疑問が長く議論されたきた(例えば渡部・平・井上,1989;池田,1992などに概観されている)。小論文に取って代わる方法も提案されてはきたが、世間一般では小論文のテスト機能に関する「信仰」は根強い。

やや古いが、2000年度の大学の入試形式に関して私立大学325校(文系64%,理系24%)を対象に調査した結果では、70%の大学で小論文試験を入学者選抜に使っていた(柳井,2002)。同様に、2012年度の国公立大学の入学試験では、大学の約8割が小論文の試験を実施していた(新井・石岡・宮埜,2013)。これらの数字は、小論文が現実選抜のためのテスト道具の役割を担っていることを明確に示している。小論文テストでは学科

の試験ではとらえにくい総合的思考力を測定できる。小論文は受験者の創造的思考力を調べることができる。小論文は作文能力を調べることができる。このような説明を多くの人は聞いているであろうが、小論文がどのように学力測定(あるいは、能力の測定)に役立つかの科学的証拠は実は乏しく(例えば、渡部・平・井上,1989)、現段階では科学的証拠に基づいた学力(能力)測定方法とはいいがたい。特に測定論の観点からは、解決すべき問題が多くある(宇佐美,2013)。研究の進展とは無関係に現実の需要にから、小論文を使って受験者の学力のある側面を測定しようとする趨勢は衰えていない。このような現状において、心理測定法(Psychometrics)にかかわってきた心理学者の責務は、小論文テスト法の限界を説明することよりもむしろ、現行の小論文テストが可能な機能を果たすことを、自分たちの知識と技能の範囲で保障することにあるかもしれない。小論文テストは、選抜のための試験道具として使いやすい道具はいえない。測定対象(能力)を限定しにく

\*岩手大学教育学部

く、また評価の数量化も多分に恣意的である。すなわちテストとして妥当性と信頼性とは常に疑問である。とはいえ、実際に使われているのであるから、わかっている弊害を取り除く努力は必要であろう。

テストを実施するにあたっては、作問と評価(採点)の両方で多くの仕事が心理学者に残されている。ある能力をいかにして測定するか。たとえば、創造性、現実の問題を把握する能力、問題を解決する能力、論理的な思考過程を適切に伝える能力などは、どれも人間の重要な知的能力の一部であり、これを測定することは実用的価値がある。テストの目的を達成できるようにテストを設計する作業は、重要で困難な仕事である。「特性Xを、個人間の固有な変動を表すものと仮定し、この特性Xを測定するように見えるテストが、事実上も、それを測定しているのだと仮定するだけでは、実はじゅうぶんではない」(Guilford, 1954)ことをテストに関わる心理学者は知っている。実際に目指している特性Xを測定しているという証拠を集めることもテスト実施者の責任、とくにテストの説明責任(accountability)に含まれる。もう一方には、採点がある。どのように答案を採点し、得点をどのように集計するかにかかわる。この側面は、作問に比べると些細で働く余地が小さいように見えるが、実際は解決されるべき問題と方法の発展の可能性がある。適切な採点は、作問と同じく適切なテストの実施に欠くことができない。

小論文の採点に関しては、心理学者に長年よく知られている方法を採点に応用する余地はある。本論文の目的は、このよく知られている知識と技術を教育実践に適用することで、試験の公平性が高まることを示すことである。ここに述べる方法を実際に採用するかどうかは、試験の管理と実施にあたる人たちの仕事であろう。

### 分散調整済み得点

採点者が報告する得点を素点とよぶことにする。同一答案に対して複数の採点者から報告された素点は、多くの場合そのまま単純に平均を計算

して(小数点第1位を四捨五入した後)得点表に記載されると想像できる。手続きが簡単であれば誤りが入りにくいし、多くの人にその手順が容易に理解されるという利点がある。このような理由から、素点の集計に単純な平均を使用することは理に適っている。しかしながら、平均のような簡明な手順が受験者間に不公平を招く場合は、その不公平を取り除く手続きを採用することは、試験実施者(あるいは、採点責任者)が考慮すべき事項となろう。

小論文採点における不公平の源泉の1つは、通常の採点過程においては、採点者個人が採点した素点の分布の大小である。工夫して採点基準を設定しても、実際の採点では個人の「くせ」が現れる。例えば、作文が得意ではない採点者は、目標とする平均点に近いところに素点が多く分布する傾向がある(中心化傾向)。一方で、採点に慣れた採点者は素点に大きなバラツキを与えることができる。平均点を意識してそれを達成することは可能であるが、特定の分散を意識してそれを達成する採点は困難である。そのため、採点者間の素点の分布は統制されずに残る。特に分散の大きさが重要である。なぜなら採点者間で分散が異なると平均は、採点者の評価を等分に加算したものとはならないからである。後述するように分散が大きい採点者の評価は、分散が小さい採点者の評価を上回る影響力を持つ。

通常入学試験のような大きな試験で小論文を採点する際に、複数の採点者を使う理由は、1つの答案に対する異なる評価を対等に勘案して最終評価を導くことにある。測定一般と同じく1つの素点には、真の値と誤差が含まれていると考えることができる。したがって複数の素点を平均するならば、誤差が独立であるという仮定のもとで、その誤差は相殺されより真の値に近くなる。これが、複数の採点者を使う基礎にある考えである。以下に説明するように、採点者の素点の分散が異なると、この目的が達成されないため、分散を調整することが望ましい。なお、得点の分散は実際場面では、その平方根(すなわち標準偏差)を使って

検討される。素点の分散を調整するとは、実際は、その標準偏差を等しくすることである。

以上をまとめると、複数の採点者による素点を集計する際には、素点の平均を計算する前に得点の標準偏差を等しく（分散を調整）するならば、各採点者の意見が等分に最終得点に反映されるといえる。本論文では以下に、まず採点者間で素点の標準偏差の大きさが異なると、これがどのように最終的な受験者の順位に影響するかを、具体的な数値例によって説明する。その後、採点者間で素点の分散が異なる場合、その違いの大きさが平均に及ぼす影響を推測する。最後に、採点者間で素点の分散が異なる場合、素点の平均がもたらす不公平の程度を受験者の順位の変化から推測する。

**数値シミュレーションによる例証**

(1) 素点の標準偏差の大きさがどう順位に影響するか――直感的に理解できる例示

現実的ではないが、採点者が2人の場合を考える。同一の答案を採点する2人の素点の標準偏差がどのように影響するか、さらにそれを是正する方法を例示する。

100人の受験者の答案を2人の採点者（AとB）が独立に採点したとする。素点は、0から100の範囲とし、平均はともに50、素点の標準偏差はAは10、Bは20とする。2人の素点には関連がないとする（相関がない）。このような制約のもとで正規分布に沿った乱数を発生させて、仮想的な素点を作成した。

2人の素点と平均の合計3種類の得点があり、それに対応して3通りの得点順位がある。どちらかの採点者（AかB）の素点が、平均点による順位に強い影響を与える時に、その影響を視覚化するために、2次元のグラフで1つの軸に平均点による順位を、他方の軸に1人の採点者の素点の順位を配して、2つの値（平均値の順位と素点の順位）の関連を図示した。

縦軸には2人の平均にもとづく順位を、横軸にはAの順位 (Fig.1a) とBの順位 (Fig.1b) を別々

に選び2つのプロットを作成した。図中の点は、受験者個人を表す。図を見やすくするために、最低得点者を1番、最高得点者を100番としているが、実際はこれはどちらでもよい。データ点は受験者個人を表す。

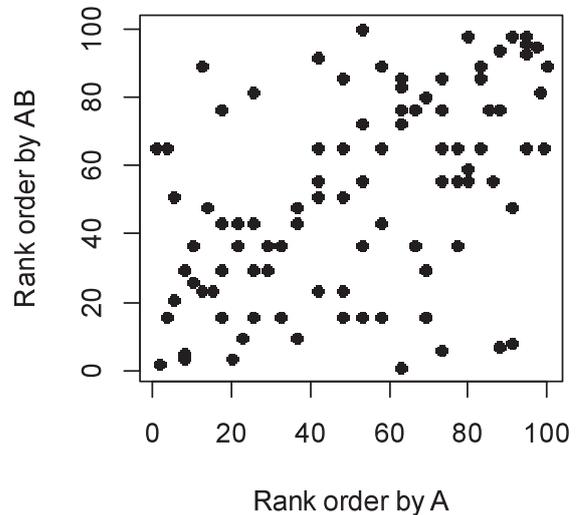


Fig.1a Aによる順位とABによる順位

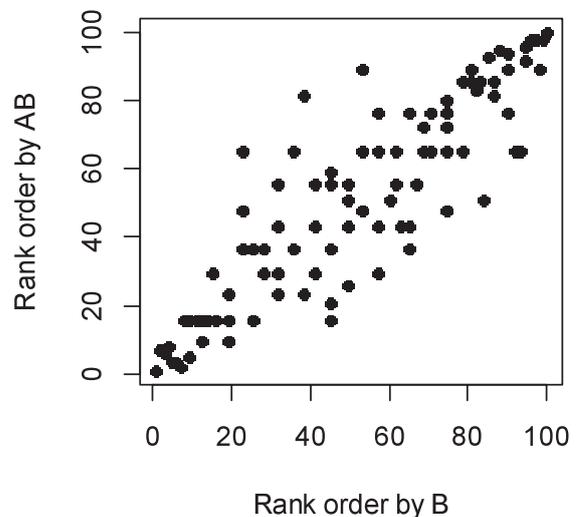


Fig.1b Bによる順位とABによる順位

図の見方は、2変数の相関を表した散布図の見方に似ている。関連が強ければ、図の右上がりの対角方向にデータの点が並び、関連が弱ければデータの点は全体に一様に分布する。Aの素点による順位とABの平均による順位の間は、やや右上がりの対角方向に点が分布する傾向がうかが

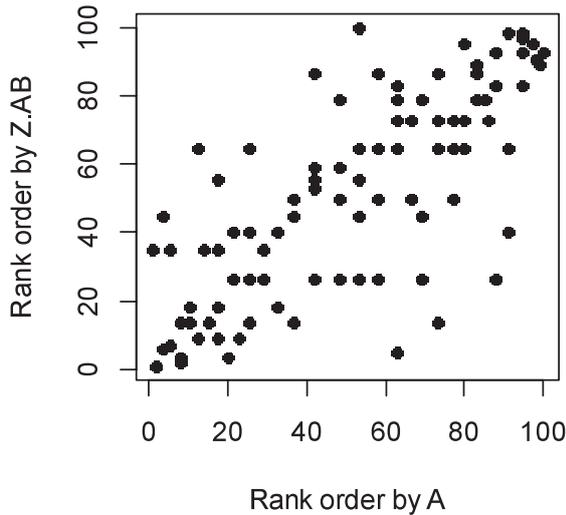


Fig.2a Aによる順位とZ,ABによる順位

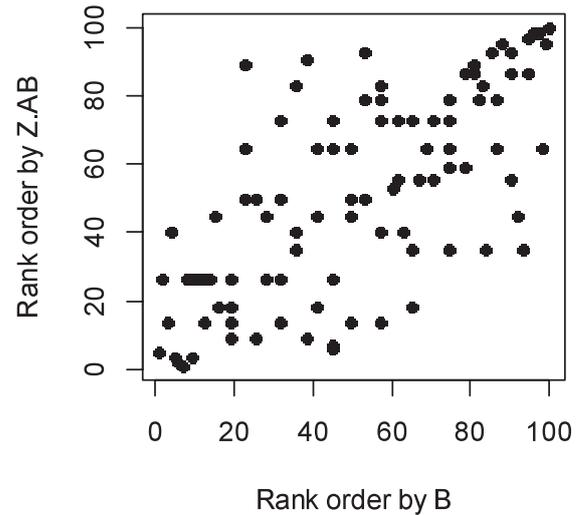


Fig.1b Bによる順位とZ,ABによる順位

えるが明確ではない (Fig.1a)。一方、Bの素点の順位とAB平均の順位の間には、明らかな関連が見える (Fig.1b)。Bの順位が上位になると平均の順位も上位になる傾向がある。2変数の順位の間を数値でみるため、順位相関係数 (Spearmanの順位相関係数) を調べると、Aの順位とABの順位の間は0.49、Bの順位とABの順位の間は0.88であった。相関係数からもBが平均の順位に強い影響を与えていることは明らかである。Fig.1.abの比較と順位相関係数の比較から、素点の平均にもとづく順位は、標準偏差が大きいBの採点者の得点順位を大きく反映していることが理解できる。

次に、同じデータを使って、得点の分散を調整するために、A、Bの素点をそれぞれ平均0、標準偏差1の値 (一般にz値とよばれる) に変換後、その平均を求め、さらに平均50、標準偏差10となる得点を作成した (一般に偏差値あるいは大文字のZを用いてZ得点とよばれる)。分布の型には修正を加えなかった。さらに通常のテスト得点に合わせるために少数点第一位を四捨五入して、整数の得点にした。順位相関を調べると、Aの順位と分散調整済みABの順位の間は0.73、Bの順位と分散調整済みABの順位の間は0.69であった。A,Bどちらかに偏した順位の間はなく、A,Bの順位が公平に反映されている。A,Bの

順位と分散調整済みABの順位を Fig.2a,b に示した。上述の手続きで、重要なのは素点の分散を調整するために使われたz変換である。このz変換により平均がAとBの評価が等しく反映する条件を作ることができた。Z得点は、z変換による得点の数値を通常のテスト得点と同じように読みやすくするための変換に過ぎない。

## (2) 分散の影響の推測

素点の分散の違いは、どの程度順位に影響するのであろうか。この疑問に答えるために数値シミュレーションを行った。標準偏差の大きさが、どのように平均値を使った順位に影響するかを調べるために、Aの標準偏差を5に固定し、Bの標準偏差を5から20まで段階的に変化させた。平均はともに50とした。このような制約のもとで正規分布に沿った乱数を発生させて、AとBの素点を作成した。AとBの素点にもとづいて、それぞれの順位と平均による順位の間を調べるために、Bの標準偏差の各段階ごとに100回のシミュレーションを行い順位相関係数の平均を計算した。それぞれの条件の値 (グラフ上の点) は100回の数値シミュレーションの平均である。

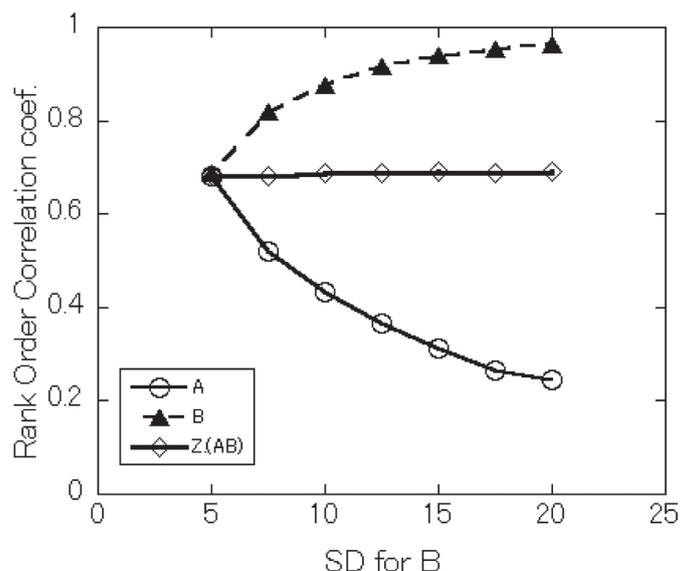


Fig.3

AとBの平均値の順位とA(あるいはB)の順位的相关係数をBの標準偏差の関数として表した。

縦軸はAとBの平均値にもとづく順位とAの素点の順位(およびBの素点の順位)の順位相関係数である。横軸はBの標準偏差である。Bが平均点の順位に及ぼす影響(▲)は、その標準偏差の増大にともない、単調に増加した。これはBの相関係数の上昇曲線に現れている。他方、Aが平均点の順位に及ぼす影響(○)は逆にBの標準偏差の増大にともない、単調に減少した。これはAの相関係数の下降曲線に現れている。図中のZ(AB)はAとBの分散を調整後に平均を作成した場合を表す。この図でもz変換を使って分散を調整した後に平均を計算するとA、Bの分散から独立した平均を計算できることが示されている。

素点の分散の違いの影響を数値でみるために、AとBの得点がAとBの平均による順位に及ぼす影響の大きさを、AとBの順位相関の比(Bの順位とAB平均の順位的相关/Aの順位とAB平均の順位的相关)で表し、これをAとBの標準偏差の比(Bの標準偏差/Aの標準偏差)の関数として表現すると、実は傾き1の一次関数となる。言いかえ、AとBの平均による順位は、AとBの得点の標準偏差の比に比例して標準偏差の大きい採点者の影響を受けるのである。

以上の説明から、採点者の得点の分散の違いが

どのように平均得点の順位に影響するか明らかであろう。2人の採点者の評価を公平に取り入れているように見える素点の平均は、実は極めて不公平な評価となりえる。その理由は、平均が公平であるのは、平均に使われる素点が等しい分散をもつという仮定が満たされた場合に限られるからである。

### (3) 順位の変化

これまで、採点者の評価を公平に集約するという側面を中心に問題を検討してきた。問題のもう一方には受験者に対する処遇がある。試験には受験者がいる。試験の公平性とは、受験者が公平な処遇を受けるということである。得点の集計に関して具体的に受験者の公平な処遇を見ると、それは個人の順位が素点の平均と分散調整済み平均ではどの程度異なるのが問題となる。

この影響の大きさを推定するために、現実的な場合を想定した。100人の受験者が問題に回答した場合を考えた。その答案を3人の採点者が独立に採点した。採点者の素点の間にはおよそ0.25の相関(ピアソンの積率相関係数rが0.05~0.40の間)があった。素点の平均は3人とも50であったが、標準偏差はおよそ10, 15, 20という具合に異なっていた。このような条件を満たすように乱数

を発生させ、相関をもたせるために素点の一部を採点者間で同じにする数値の入れ替えによって3人の素点を作成した。調べたいのは、調整済み平均に基づく受験者の順位が素点の平均に基づく受

験者の順位によってどの程度変化するかである。結果を図で示す前に、単純な例を図解する。素点の平均と調整済み平均という2種類の平均による順位を、3人の受験者を想定して図示した。

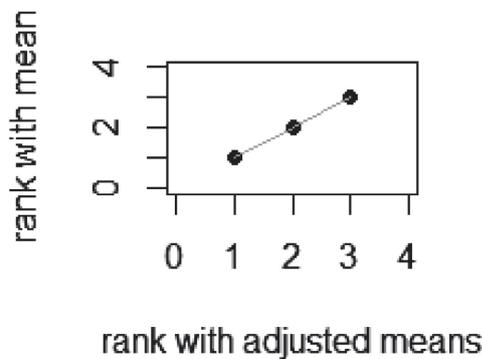


Fig.4a

横軸は調整済み平均による順位、縦軸は素点の平均による順位を表す。素点の平均による順位と調整済み平均の順位が一致する場合を示した。このような事態では、受験者の順位が素点の平均で決まっても、調整済み平均で決まっても違いはない。上述したように、素点の分散が採点者間で等しければ、この図で示すような結果となる。すなわち、素点の平均が公平な結果につながる。

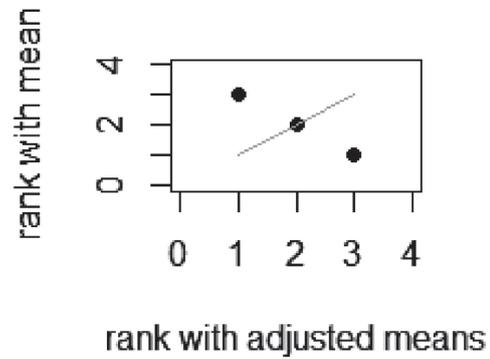


Fig.4b

横軸は調整済み平均による順位、縦軸は素点の平均による順位を表す。素点の平均による順位と調整済み平均が一致しない場合を示した。調整済み平均では1位（左上の黒丸）の人が、素点の平均では3位に下降している。逆に調整済み平均で3位の人（右下の黒丸）が、素点平均では1位に上昇している。このような事態では、素点の平均による順位は公平な結果につながらない。

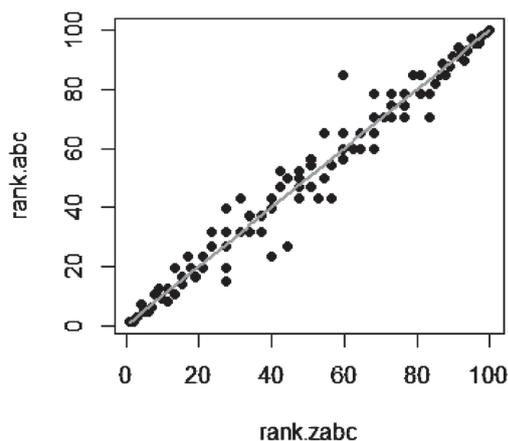


Fig.5

シミュレーションの結果を示す。横軸は調整済み平均による順位、縦軸は素点の平均による順位を表す。右上がりの対角線から外れた点は、全て順位が変動した個人である。

Table 1. 3人の採点者 (a,b,c) の素点の相関行列

	a	b	c
a	1.00	0.21	0.38
b	0.21	1.00	0.28
c	0.38	0.28	1.00

Table 2. 3人の採点者 (a,b,c) の素点の標準偏差

	標準偏差
a	10.49
b	14.42
c	21.11

Table 3 分散調整済み平均から単純平均に変えた場合の順位の変化

順位	平均	標準偏差	A,B,Cの特性	平均	標準偏差
1位下降	8.29	3.19	A,B相関	0.05	0.11
2位下降	7.00	2.77	B,C相関	0.24	0.10
3位下降	5.98	2.62	A,C相関	0.35	0.08
4位以上下降	26.95	4.10	Aの標準偏差	9.47	1.17
変化なし	10.80	3.48	Bの標準偏差	14.45	1.15
上昇	40.97	3.37	Cの標準偏差	20.43	1.31

Fig.5にシミュレーションによる仮想データを示した。Fig.5では上位の10位以内から100位に近い下位まで、多くの位置で順位の入替わりが見える。右上がりの対角線より上に位置する点は、素点の平均に基づいて順位を決めると順位が低下する人である。特に目に付くのは横軸で60付近で縦軸で80~90にある人であるが、この人は、調整済み平均では60位前後であるが、素点の平均では80位以下の順位となる。幸い多くの点には、これほど極端に大きな順位の変動はない。

シミュレーションで設定した条件は現実の小論文採点の条件に近い(阿久津ほか, 2005; 阿久津ほか, 2006)ため、Fig.5に示したような結果は現実に生じる。この図からみると、大きく順位が変動する個人は多くはないが、それでも右上がりの対角線から上に位置する個人を不利益を受ける順位変動者とする、受験者の10%にあたる10人程度の受験者に3位以上の不利益変動が生じている。

ここで設定した条件でどの程度の順位の変動がおきるかを推測するために同じ条件で300回のシミュレーションを行った結果をTable 3に示した。Table 3の左端の順位に対応する右の数字は、該当する人数を表す。順位が3位下降する人と4位以上下降する人を合わせると約32人であり、全体の3分の1に近い受験者に実質的な順位の変化が生じた。もちろん反対に順位が上昇する受験者もいる。順位に変化がない受験者は、全体の約10%に相当する11人程度であった。言い換えれば、素点の平均を使うことによって、およそ90%の受験生が不当な順位を与えられる。もっとも、不利益

の評価は難しく、例えば最上位近くに位置する受験者にとっては、3位か4位の順位下降は、実質的には試験の結果(合格と不合格)に影響せず不利益とは言えないであろう。一方では、大学入試センターの学力試験では、試験の実施時間を2~3分短くして終了すると、それは受験生に対する不利益であり、試験実施の不祥事と判断される。この不利益と比較して、Fig.5にある採点の際に生じる受験生の不利益をどう評価するかは難しい問題である。

#### 残された問題

本論文では、調整済み平均を計算する際の標準偏差に乗ずる定数に、Z得点を計算する際の慣例にしたがって10を使用した。実際に調整済み平均を求める際には、慎重にこの定数を決める必要がある。その理由は、もし他の試験(例えば面接)の結果があり、それを総合して受験者個人の最終得点を決めるとするならば、この係数の大きさは他の試験結果と合計する場合の相対的重みに影響するからである。得点の実際の分散が、その相対的重みを決定することは、Fig.3から明らかである。より正確に述べれば、標準偏差の比が、相対的重みを決定する。したがって、最高得点や得点範囲を設定するだけでは、その得点が高得点と合計されるとき相対的重みをコントロールしたことにはならない。例えば、特性Xを50点満点、特性Yを100点満点として合計得点を計算しても、特性Xに対して1、特性Yに対して2の重み付けで総合得点を計算したことにはならない。もし特性Xと特性Yを1:2の重みで評価したものを

最終得点とすると定義したいならば、特性 X の得点と特性 Y の得点の標準偏差を 1:2 とする必要がある。

### 今後の展望

これまで小論文の採点方法について述べてきたが、分散調整済み平均の使用は、テストの信頼性を引き上げるのに貢献するが、小論文試験全体を見たときには実は小手先の技術にしかすぎない。適切な小論文試験を行うには、まず妥当性がある試験問題を作ることが最も重要である（池田, 1992）が、試験問題作成は本論文の考察範囲外である。

小論文テストの作成方法や採点方法については、いくつか洗練された方法が提案されているが、おそらく現実の大学入試の準備と実施状況には、テストの専門家が提案する方法が入りこむ余裕がないと想像できる。それは、おもに実施上の問題——新しいシステムを実現するための時間と人の余裕がないこと——に起因している。また、一部にはテストに対して実施主体（大学とその教員）に正しい知識と理解が普及しにくいこともあるかもしれない。どこに原因があるにせよ、テストの専門家の知識・技術とテストを実施する機関（大学など）の実践の乖離は、今後大きくなることはあっても小さくなることはなさそうである。そのことによる不利益は受験者だけにとどまらず、大学にも及ぶであろう。そうなる現実的な解決方向は、特定の機関（例えば大学入試センター）にテストの実施を任せることであろうか。

### Appendix 1

#### 分散調整済み得点の計算方法

複数の採点者による素点を集計する際には、平均を得る前に素点の分散を調整することが適当である。実際には尺度構成法において一般に  $z$  変換とよばれている変数変換（池田, 1973, p.40）を適用する。以下にこの方法を具体的に説明する。 $x$  と  $z$  はベクトルである。

$z$  変換は次式であらわされる。

$$z_i = (x_i - \text{mean}(x)) / \text{sd}(x) \dots\dots \text{eq.1}$$

ここで  $x$  は  $x_1, x_2, x_3, \dots, x_n$  の要素からなる素点、 $\text{mean}(x)$  は素点の平均、 $\text{sd}(x)$  は得点の標準偏差をあらわす。 $x_i$  は素点  $x$  の  $i$  番目の要素である。 $z_i$  は変換された  $z$  の  $i$  番目の要素である。

$z$  は平均 0、標準偏差 1 の分布の特性をもつ。

各採点者が報告した素点を  $z$  値に変換し、その平均を求める。

もし設問が 2 つある場合は、2 つの設問について、同様な処置を行う。

最後に、小論文試験の得点を、受験者ごとに作成する段階となる。2 つの設問については予め決められた重みがある。重みの合計を 1.0 として、設問別に受験者の得点（ $z$  の平均）に重み係数を乗じ、設問別の得点を加算する。設問 1 の重みを  $w_1$ 、設問 2 の重みを  $w_2$  とすると、次式 eq.2 によって、重みつき平均を求める。

$$z_{\text{総合}} = z_{\text{平均}} \times w_1 + z_{\text{平均}} \times w_2 \dots\dots \text{eq.2}$$

ここで  $x$  は乗算をあらわす。

得られた  $z_{\text{総合}}$  は、平均は 0 であるが、標準偏差は 1 ではない。そこで、再度、eq.1 を  $z_{\text{総合}}$  に適用して、 $z$  値を求める。最後に期待する平均（例えば 50）と採点過程で得られた得点の標準偏差およびその他の要因を勘案した標準偏差を使い、次式 eq.3 により得点を求める。

$$\text{得点} = \text{期待平均} + z \times \text{標準偏差} \dots\dots \text{eq.3}$$

この得点（小数点第 1 位を四捨五入した整数）が、最終的報告書に試験の得点として記載される。

### Appendix 2

#### 仮想的データを使った方法の説明

同一の問題に対して 100 人の受験者が回答し、その回答を 3 人の採点者が採点する事態を仮定した。素点と分散調整済み得点の計算方法を以下に述べる。統計ソフトの R を使って計算した（cran.r-project.org）。エクセルの場合は、素点の列の最下段（最後の行）の 2 つに、平均と標準偏差を書き、それらの数値（セルの絶対番地を参照して）を利用して素点の右の列（4～6 列目）に  $z$  値を書き

入れるように指定する。その右にzの平均、さらに右に平均値と標準偏差を指定した最終得点を書

き入れるように指定すればよい。Rを使えばそのような煩雑な数値計算は数行の式で実行できる。

Table 1. 100人の受験者に対する3人の採点者による素点

s	a	b	c
s1	48	40	42
s2	72	47	45
s3	57	57	69
s4	38	40	50
s5	50	51	47
s6	38	49	48
s7	44	42	48
s8	57	56	74
s9	68	57	58
s10	52	44	26
s11	79	46	44
s12	64	48	48
s13	40	54	53
s14	46	55	63
s15	50	51	76
s16	53	48	50
s17	50	46	44
s18	49	52	64
s19	46	57	40
s20	42	45	55
s90	37	49	25
s91	64	59	55
s92	58	50	61
s93	59	50	56
s94	49	49	46
s95	56	50	27
s96	42	44	7
s97	67	48	79
s98	25	46	44
s99	38	53	58
s100	58	48	27

Table 2. 採点者の素点の平均と標準偏差

	a	b	c
mean	49.44	49.52	50.84
SD	10.78	4.85	13.88

Table 3. 採点者間の素点の相関

	a	b	c
a	1.000	0.051	0.110
b	0.051	1.000	0.165
c	0.110	0.165	1.000

数値計算表のエクセルを使っても、分散調整済み得点を計算できるが、汎用統計ソフトRを使うと簡単である。Rを使って分散調整済み得点を求める方法を以下に述べる。

1. まずエクセルファイルで、a,b,cの素点をa,b,cのラベルも含めてRのworking spaceに読み込む。

```
d<- read.delim('clipboard')
```

head(d) #内容を画面に提示して目で確かめるため

2. 計算

dに含まれる変数に簡単にアクセスするためにattach(d)

次いで、以下のコマンドを実行してzを求める。

```
za<-(a-mean(a))/sd(a)
```

```
zb<-(b-mean(b))/sd(b)
```

```
zc<-(c-mean(c))/sd(c)
```

```
z<-(za+zb+zc)/3
```

```
z<-(z-mean(z))/sd(z)
```

```
zz<-round((z*20+50))
```

この最後のzzが分散調整済み得点である。

あらかじめ自作関数として

```
myz<-function(d){
```

```
  a<-d[,1]
```

```
  b<-d[,2]
```

```
  c<-d[,3]
```

```
  za<-(a-mean(a))/sd(a)
```

```
  zb<-(b-mean(b))/sd(b)
```

```
  zc<-(c-mean(c))/sd(c)
```

```
  z<-(za+zb+zc)/3
```

```
  z<-(z-mean(z))/sd(z)
```

```
  zz<-round((z*20+60))
```

```
  return(zz)
```

```
}
```

を作成して、このソースコードをRに読み込んでおけば、

```
d<- read.delim('clipboard')
```

```
z<-myz(d)
```

を実行すると、zに調整済み得点を得ることができる。

## 引用文献

- 阿久津洋巳, 嶋野美恵子, 熊谷賢, 佐々木和歌  
(2005) 課題レポート評価における評定者間  
の一致 岩手大学教育学部附属教育実践総合  
センター研究紀要 第4号, 75-80.
- 阿久津洋巳, 菊地梢, 鈴木安澄, 鈴木光, 渡邊愛  
枝 (2006) 論述式テストの研究(1): 採  
点者間の一致度 岩手大学教育学部附属教  
育実践総合センター研究紀要 第5号, 115-  
122.
- 新井清佳, 石岡恒憲, 宮埜壽夫 (2013) 大学入学  
者選抜における小論文試験と教科・科目試験  
との関連について 日本テスト学会誌 第9  
巻, 28-36.
- Guilford, J.P., (1954) *Psychometric Methods*,  
McGraw-Hill Book Company, Inc. New York.  
(日本語訳 J.P. ギルフォード (1959) 精神  
測定法 秋重義治 監訳 培風館 p.522)
- 池田央 (1992) テストの科学 日本文化科学社
- 池田央 (1973) 心理学研究法8 テストII 東京  
大学出版会
- 池田央, 芝祐順 (1972) テスト法の意義 心理学  
研究法7 テストI 東京大学出版会
- The Comprehensive R Archive Network [https://cran  
r-project.org](https://cran.r-project.org)
- 宇佐美慧 (2013) 論述式テストの運用における測  
定論的問題とその対処 日本テスト学会誌  
第9巻, 145-164.
- 柳井晴夫 (研究代表者) (2002) 大学入学者選抜  
資料としての総合試験の開発的研究 研究課  
題番号 11551004 科学研究費補助金 研究成  
果報告書
- 渡部洋, 平由美子, 井上俊哉 (1989) 小論文評価デー  
タの解析 東京大学教育学部紀要28巻, 143-  
164.