高等学校数学科「数学Ⅰ」における統計指導での変数変換の扱いに関する実践研究

半田 真 [a)]　清水 克彦 [b)]

要旨：高等学校数学科「数学Ⅰ：データの相関」の指導において、変数変換 Y=aX+b の指導のあり方を提案する。変数変換 Y=aX+b を用いた平均、分散等を表計算ソフト Excel を用いて求める中で成り立つ性質を生徒に気づかせ、その理由を証明する意欲を引き出す指導を試みた。その結果、変換した変量 Y について、表計算ソフトで統計量を求めることは 108 名中 49 名、約 45% の生徒ができていたことが確認できた。一方、$E(Y)=aE(\mathrm{X})+b$、$V(Y)=a^2V(X)$ が成り立っていることに気づく者はいなかった。生徒は直観的には理解しているものの、それを数式を用いて証明するまでには至っていなかったことが確認できた。

キーワード：統計指導、表計算ソフト、変数変換、証明、意欲

Ⅰ．研究意図と目的ならびに研究計画

2009 年告示、2012 年施行の高等学校学習指導要領（以下、現行学習指導要領）では、数学Ⅰで統計分野の指導が必修となっている。それにともない大学入試センター試験でも数学Ⅰで統計分野の出題が 2015 年以降必ず出題されている。

本研究では、高等学校数学科「数学Ⅰ」の統計指導において、学習指導要領に示されているように ICT の活用として表計算ソフト Excel の活用を前提とした指導を行うこととした。Y=aX+b などの変数変換、及びそれを用いた平均、分散等の指導に関して、生徒が表計算ソフト Excel を用いてデータを操作する中から成り立つ性質（公式）を生徒に気づかせ、それが成り立つ理由を証明しようと生徒自ら考えられるようになることを検証していくことを目的とした。

具体的にはまず、教材開発に際し、現行学習指導要領のもと 2015 〜 2018 年に行われたセンター試験数学 IA「データの分析」における出題傾向を確認した。特に変数変換、及びそれを用いた平均、分散等の計算に関する出題に注目し、教材開発の対象とした。

それに従い本稿では数学Ⅰ「データの分析」で確率変数の変数変換に関する指導法及び教材を提案する。提案した教材を用いて都内中高一貫の私立校普通科で授業実践を行う。

Y=aX+b などの変数変換、及びそれを用いた平均、分散等の計算を数学Ⅰ「データの分析」でも指導する必要性を確認する。大学入試センター試験問題評価委員会報告書（2017、p.166）によると、2017 年度のセンター試験の問題作成部会の見解では、数学Ⅰデータの分析の出題における変数変換の扱いには「変換に関する基本的内容を問う問題である。」とある。分散や共分散・相関係数の定義を単に暗記するのではなくその意味を理解した上での計算を求めているものと思われる。そこで本研究においても数学Ⅰ「データの分析」で確率変数の変数変換を行い、変換前と後で分散や共分散・相関係数が変化するのか否か確認できるような教材を開発、授業実践を行う。その教材による指導で、変数変換及びその計算等に関する学

[a)] 理学部第二部 物理学科　[b)] 理学部第一部 数学科

習効果があったことを教材ワークシートにおける学習状況や事前･事後の検証問題で確認することとした。

Ⅱ．現行学習指導要領「数学Ⅰ」における統計指導と教材開発の指針

2009 年公示の高等学校学習指導要領解説では数学Ⅰ「データの分析」の単元目標として「統計の基本的な考えを理解するとともに、それを用いてデータを整理・分析し傾向を把握できるようにする。」（文部科学省、2009、pp.24-25）とある。データの相関に関しては「イ．データの相関散布図や相関係数の意味を理解し、それらを用いて二つのデータの相関を把握し説明すること。」（文部科学省、2009、pp.24-25）とも記されている。一方で、Y=aX+b の変数変換の扱いについては触れられていないが、分散や共分散・相関係数の定義を単なる暗記ではなく、理解させることで「基本的な考えを理解する」（文部科学省、2009、pp.24-25）の中に変数変換した場合の分散や共分散・相関係数をどのように求めればよいか考察できることを含んでいるものと筆者は捉えた。

1．センター試験「数学Ⅰ」での統計分野の扱い

現行学習指導要領対応のセンター試験は 2015 年から実施され今年で 4 年目である。その数学 IA 本試験において、次のような 2 つの傾向があると半田（2017）は指摘する。一つ目は「センター試験では「数学Ⅰ」の各社検定教科書ではさほど強調されていない変数変換 Y=aX+b 及びそれらを用いた平均、分散、標準偏差に関する出題があることが分かった。」（半田、2017、p.72）である。確率変数の変数変換は数学 B の「確率分布」で扱う検定教科書が多いが、センター試験では図 1 に示すような問題が 2017 年に出題されている。同様の問は、2016 年にも見られた。

二つ目は「統計量の定義や意味などを理解した上で､求めた統計量やグラフ等が何を表しているか､デー

(2) 得点 X は，飛距離 D から次の計算式によって算出される。

$$X = 1.80 \times (D - 125.0) + 60.0$$

次の $\boxed{ソ}$，$\boxed{タ}$，$\boxed{チ}$ にそれぞれ当てはまるものを，下の⓪～⑥のうちから一つずつ選べ。ただし，同じものを繰り返し選んでもよい。

- X の分散は，D の分散の $\boxed{ソ}$ 倍になる。
- X と Y の共分散は，D と Y の共分散の $\boxed{タ}$ 倍である。ただし，共分散は，2つの変量のそれぞれにおいて平均値からの偏差を求め，偏差の積の平均値として定義される。
- X と Y の相関係数は，D と Y の相関係数の $\boxed{チ}$ 倍である。

⓪ -125	① -1.80	② 1	③ 1.80
④ 3.24	⑤ 3.60	⑥ 60.0	

図 1　2017 センター試験　数学Ⅰ・数学 A より

タから何が読み取れるかを聞く問が明らかに増えている。」（半田、2017、pp.72-73）である。
本研究は、一つ目の変数変換に関する扱いとその指導に関する実践研究である。

2. 変数変換とその計算の指導の必要性

学習指導要領では数学 I 「データの分析」において「統計の基本的な考えを理解する」（文部科学省、2009、pp.24-25）の中に変数変換を含んでいるものと解釈する。さらにセンター試験数学 IA では変数変換 Y=aX+b 及びそれらを用いた平均、分散、標準偏差に対する指導が求められている記述もあることから、数学科の統計指導でも数学 I の段階で変数変換 Y=aX+b 及びそれらを用いた平均、分散、標準偏差の指導が求められていると言えよう。その際、公式として以下が

$$E(Y)=aE(X)+b,\qquad V(Y)=a^2V(X)$$

が成り立つことを伝えるだけではなく、定義に基づいた理解が欠かせない。そこで、本研究では実際に多くのデータを操作し、上記の式が成り立つことに気づかせるような指導を取り入れていくこととした。さらにこれらの公式がなぜ成り立つのかを考えさせ、証明しようという意欲を引き出すような指導を目指す。この指導により単なる公式の暗記ではなく定義に基づいた理解ができるようになると思われる。

3. 教材開発に向けた検討

ICT の数学科での利用について、清水・垣花（1999）は「なんらかの『予想』や『観察したい性質』があるときには、このように具体例を通した検討が有効である」（清水・垣花、1999、p.16）として動的幾何ソフトを思考実験のための実験装置として利用することを強調している。本研究でも表計算ソフト Excel の利用の有効性を取り入れることとした。具体的には本研究は統計指導に表計算ソフト Excel を実験装置として利用し、データから一般的な性質に気づかせる指導を取り入れることとした。そうした指導を通じて「生徒が発見的・探求的な活動を行い、自ら知識を発見するという活動のなかで、生徒自ら『いつでも成り立つ』知識を見いだすことで、証明への意欲や、証明を書くときの根拠を得ることを支援」（清水・垣花、1999、p.19）することが表計算ソフト Excel を利用するメリットのひとつである。

また、佐伯（2010）は「生徒の身の回りにある事象に対して、生徒が興味・関心をもって事象と深く関われるような仕掛けを教材にもたせるべきだ」（佐伯、2010、p.190）と指摘する。そこで、本研究では扱う統計データを実践校の過去の生徒の体力測定データを用いることとした。このような、身近なデータを用いることで教材に生徒が興味・関心をもてるようにした。

Ⅲ. 授業実践の様子

本章で提案する教材を用いて、数学 I 「データの相関」を学習した。対象は、都内の普通科私立高等学校一年生である。表計算ソフト Excel を用いるため、情報科の授業として実践を行った。

1. 対象とする生徒の実践前の学習状態

対象生徒は数学 I 「データの分析」の単元を 1 学期に数学 I の授業で学習済みである。しかし、その授業は座学のみのため表計算ソフト Excel を全く用いない形で行われた。

一方、生徒は表計算ソフト Excel 及びフリーソフト GeoGebra の基本的な操作を 9 月までに情報科の授業で習得済みである。表計算ソフトの関数の扱いやグラフ描画、GeoGebra で箱ひげ図などのグラフ描画等も習得している。

授業は、週 1 時間コンピューター教室で 3 ～ 4 人のグループで課題を考える形態である。この授業実践は、合計 3 時間で実施した。

2. 実践前の診断的評価と実践後の総括的評価の実施

3学期最初の授業でⅢ章3節の教材を実践する前に、1学期数学Ⅰで学習した「データの相関」に関する小テストを実施し、実践前の診断的評価とした（問題は巻末の付録を参照)。散布図からの傾向の読み取りを聞く問いを5問、確率変数の変数変換による平均・標準偏差、共分散がどのように変化するかを聞く問いを2問（内1問は記述式）出題した。

実践前の診断的評価と実践後の総括的評価を比較することで、数学科における統計指導に本研究の表計算ソフト Excel を活用した指導が、証明すること、意欲を引き出すことに有効であるかどうかを検討することとした。

3. 実践教材とそのねらい

本研究で提案する教材が表1ある。

（1）については、2つの変量x、yの平均が和の変量 $g=x+y$ の平均になることを確認させることを目的にしている。また、分散ではその関係が成り立っていないことを確認させるのも目的とした。

（2）については、外れ値が相関係数に与える影響を確認させることを目的にしている。

（3）－（4）では、変量を一次式で変換したときに、平均・分散、共分散や相関係数がどの様に変化するか否か確認することを目的にしている。

（5）では、この文脈に沿った回答を考えさせ、クラスメイトの前で発表することで考察を深めさせることをねらいとした。

表1　本実践の教材案

A さんの学校では，高校一年生に対して入学した時点で体力測定を行います．そこでは以下の7項目が測定されています．

握力（右・左)，体前屈，幅跳び，上体起し，反復横飛，50m 走

x，y，b，j，p，q，v

こんなに多くの項目を測定するには時間も手間もかかると考えた保健体育委員長の A さんは，それぞれの項目で相関を調べ，相関の強いものは測定しなくても良いことにしたいと思いました．

そこで，過去5年間の1年生の生徒約1500人分の体力測定データをもとにそれぞれの相関を調べてみることにしました．

(1) まず，握力について左・右の相関はどうか調べました．その結果，相関係数が 0.79 と強い相関があることが分かりました. それは当然と考えた A さんは右の握力データ x と左の握力データ y の和を変量 g にすることを考えました．
右を変量 x，左を変量 y，$g=x+y$ として変量 x，y，g 平均，分散を求め，比較せよ．比較して気づいた事柄を列挙せよ．

	x	y	g
平均			
分散			

(2) 上体起し（変量 p）と反復横飛（変量 q）の相関を調べてみると，上体起しのデータ内に極端に大きな値があることに気づきます．調べてみると，23 とすべきデータを 234 と入力ミスしていたことが分かりました．
修正前の相関係数 r と修正後の相関係数 r’を比較して気づいたことを列挙せよ．

(3)　体前屈（変量 b）のデータは計測機の不具合が見つかりました．すべての元データ b を以下の式で c に変換しなくてはいけないことが分かりました．

$$c = 0.8b - 1.2$$

修正前の平均・分散と修正後の平均・分散をそれぞれ求めて比較し，気づいたことを列挙せよ．

(4)　(3)のとき，体前屈（変量 b）と幅跳び（変量 j）の相関を調べます．体前屈（変量 b）データの修正前との相関 r1 と，修正後の体前屈（変量 c）データとの相関 r2 を比較し，気づいたことを列挙せよ．また，共分散もそれぞれ求めよ．

(5)　上記の修正を行った上で A さんは次の 5 つのデータ間で相関の強いものを探してみることにしました．その結果から，測定すべき項目を減らそうと考えています．

握力 g，体前屈 b，幅跳び j，上体起し p，反復横飛 q，50m 走 v

グループごとに A さんのように相関を調べ, どの測定項目を減らしたらよいか提案せよ．（グループでプレゼンにまとめ発表できるようにすること．）

Ⅳ．授業実践からの考察

授業実践を行った 3 クラスは都内中高一貫の私立校普通科の生徒 38 名、39 名、31 名である。これらクラスは文系・理系の偏りはなく、前年度の学力評定によって学力差がないよう学年始めにクラス編成がなされている。授業は各クラスとも同一の教員が指導にあたった。表 1 のワークシートを用いた授業は各クラスそれぞれ 3 時間（内 1 時間はプレゼン発表）で実施した。

教　材：自作教材ワークシート
単元名：「データの相関」
対象クラス：都内私立校普通科高校 1 年 3 クラス 38 名、39 名、31 名（計 108 名）
授業期間：2018 年 2 月 13 日（火）～ 3 月 1 日（木）
定期試験日時：2018 年 3 月 7 日（水）、30 分

1．ワークシートへの取り組みからの考察

表 1 のワークシートでの授業に入る前に 1 学期の数学 I で学習した「データの相関」について復習した。表計算ソフト Excel を用いた平均・分散・標準偏差の求め方や散布図・回帰直線の描き方についても学習した。分散については偏差の 2 乗の平均で求める方法を学習した後、表計算ソフト Excel の VAR.P 関数を用いた計算についても指導した。

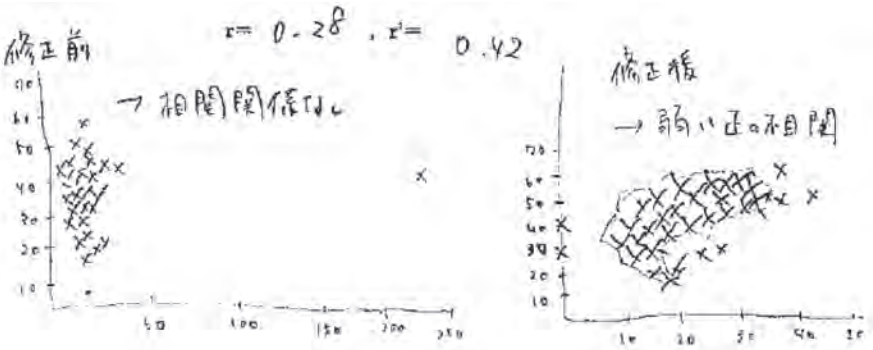


図 2　生徒のワークシート（2）から

表 1 のワークシート（1）では、Ⅲ章 3 節で述べたねらいを 108 中 95 名の生徒が確認できていた。(2) についても相関係数が 0.28 から 0.42 に変化するとともに散布図の形も変化していることを 13 名の生徒が捉えている様子（図 2）が見られた。その一方、相関係数だけ、あるいは散布図だけのどちらか一方しか記していない者が 75 名いた。そこで、教員側から「両方確認する」よう指導した。

（3）については、表計算ソフトで統計量を求めることは 49 名、約 45% の生徒ができていた。しかし、

$$E(Y)=aE(\mathrm{X})+b、\quad V(Y)=a^2V(X)$$

が成り立っていることに気づく者はいなかった。(4) についても相関係数・共分散を求めることは 28 名、約 26% の生徒ができていたが、共分散が a（=0.8）倍されていることに気づく者はひとりもいなかった。

一方、相関係数が変わっていないことについては 26 名、約 24% の生徒が気づいていた。しかし、その理由を数式で説明できた者はひとりもいなかった。相関係数が変わらないのは「それぞれの値の差が変わらない」から変化しなかったのだと指摘する者（図 3）が 1 名だけいた。分散・共分散・相関係数等の統計量の定義をある程度理解できている者が一人だけというのは教材のあり方や指導のあり方に工夫が足らなかったと思われる。

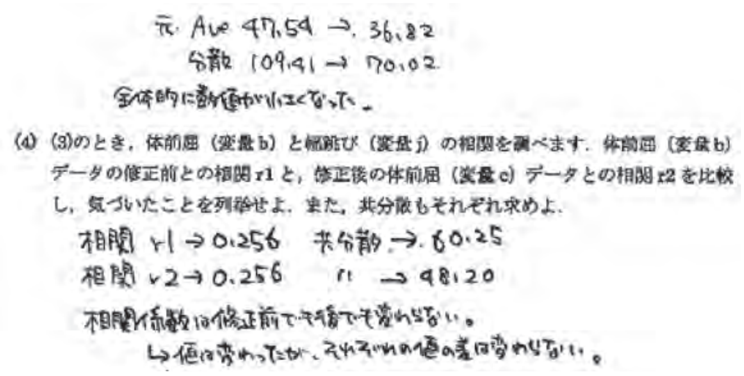


図 3　生徒のワークシート（3）－（4）から

2. 実践後の総括的評価

Ⅲ章 2 節でも述べたとおり 1 月最初の授業で事前小テストを実施した。実施後 1 週間ほどで採点したものを返却している。返却の際、問題についての解説は行ってはいない。その約 2 ヶ月後、3 月 7 日に事後テストを実践校の定期試験の一部として実施した。事後テストでは、事前小テスト（問題は巻末の付録を参照）と同じ問題を出題し、それらの問について実践後の平均点が実践前の平均点と比較して上昇したかを検討した。F 検定の結果、どの問も正解率の分布が等分散である保証が得られなかったため、等分散ではない場合の t 検定を有意水準 5% の片側検定で行った。その結果が、表 2 である。

表 2　各問の検定結果

問題番号	事前テスト		事後テスト		p-value
	正答率平均	S.D.	正答率平均	S.D.	
【1】ア	0.92	0.07	0.91	0.08	0.6732
【1】イ	0.91	0.08	0.88	0.11	0.8026
【1】ウ	0.92	0.07	0.89	0.10	0.8139
【1】エ	0.84	0.14	0.91	0.08	0.0681
【1】オ	0.80	0.16	0.83	0.14	0.2765
【1】合計	4.40	0.98	4.41	0.97	0.4673
【2】	0.31	0.22	0.55	0.25	0.0002**
【3】	0.19	0.17	0.26	0.55	0.1963
	n_1=105		n_2=107		

【1】の散布図からデータの傾向を読み取る 5 題の問については、事前・事後で平均正答率（正答者数をサンプル数で割った値）に大きな違いはなかった。なお、【1】は正誤問題のため事後テストでは小問の並びを入れ替えて出題している。【1】の結果からは、事前・事後とも 8 割以上の正答率平均である。散布図から相関を読みとることについては、本実践での影響は見られなかったと考えられる。

【2】の変数変換をした場合の平均値 E（Y）と標準偏差 V（Y）の計算については、事後の正答率平均の方が有意に高かった（p=0.0002**）。なお、【2】は選択肢から選ぶ問いのため、事後テストでは選択肢の順番を入れ替えて出題している。【2】の結果から本実践により

$$E(Y)=aE(\mathrm{X})+b、\quad V(Y)=a^2V(X)$$

について知識として習得できた可能性がある。しかし、ワークシートへの取り組み具合からは証明までしようという意欲を引き出すまでには至っていなかったことが分かる。証明まで考察させるには本実践だけでは不十分であったと考える。

【3】については記述解答を求めた問いである。事後テストでは、実践校の期末試験問題の最後の問いとして出題したため最後まで解答できていない者もいたものと考えられる。【3】の平均正答率は事前に比べて事後の方が幾分良くなってはいるが、有意な差は見られなかった（p=0.1963）。平均正答率は事前に比べ向上したものの 0.26 という平均正答率は低く、数式を用いた考察を行うには指導のあり方に工夫の余地があると考えられる。

V．まとめと今後の課題

変数変換の指導について、表計算ソフト Excel 利用を通じて生徒の気づきを引き出し、その気づきを演繹的に証明しようという意欲を引き出すことを目指した。だが、変数変換によって平均や分散がどの様に変わるのかについて具体例から

$$E(Y)=aE(\mathrm{X})+b、\quad V(Y)=a^2V(X)$$

に気づくことは難しかったようである。Y の分散を求める問については、a（=0.8）倍ではなく、2 倍とか 1/2 倍など気づきやすい数値にしておくべきであった。しかし、相関係数が変化しなかったことに気づき、データの「値の差が変わらない」から変化しないと考える生徒（図 3）がいたことがワークシートの取り組みからも確認できた。この生徒は、統計量に関する定義は理解していたものと考えられる。だが、数式を用いて理由を説明するまでには至っておらず、課題が残った。それは表 2 に示した各問の検定結果からも確認できる。選択肢から選ぶ【2】の変数変換の問に関しては実践前と後で向上したことが確認でき、知識としての習得できた生徒がいた可能性があると考えられる。しかし、証明までしようという意欲を引き出すまでには至らなかったようである。【3】のような数式を用いて説明する証明問題でも有意な差は見られなかった。今後は、数式を用いた証明に関する指導のあり方についてどうあるべきか再考が必要であると考える。

参考文献

大学入試センター試験問題評価委員会報告書（2015-2017）　http://www.dnc.ac.jp/data/hyouka.html（最終確認 :2017.6.23.）

半田真（2017）「センター試験の統計問題傾向変化と今後の展望－ 統計が必修化された現行学習指導要領における統計分野の扱い－ 」数学教育東京理科大学数学教育研究会会誌 Vol.59-no.02 ’17, pp.69-73

文部科学省（2009）「高等学校学習指導要領解説（平成 21 年 7 月）－数学編」, pp.24-26

佐伯昭彦（2010）「高等学校 新学習指導要領の展開 数学科編」吉田明史 編著 , p.189-193, 明治図書

清水克彦・垣花京子（1999）「コンピュータで支援する生徒の活動」清水克彦・垣花京子 編著 , p.15-19,

明治図書
東京理科大学数学教育研究所（2018）「2017 年度理数系高校生のための数学基礎学力調査」報告書（中間）, pp.47–60

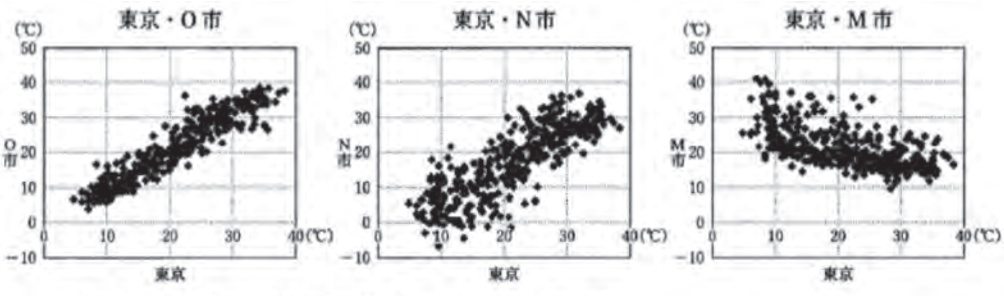
付録

評価問題

Ⅲ章 2 節及びⅣ章 2 節で実施した評価問題を挙げる。大問 3 題、小問も含め合計 7 題である。最後の問以外はすべて選択肢問題になっている。【1】は 2016 年大学入試センター試験「数学Ⅰ」の出題の一部から選んだ問いである。【2】と【3】は東京理科大学数学教育研究所（2018）が 2017 年度に実施した学力調査問題の中から選んだ 2 問である（東京理科大学数学教育研究所、2018、pp.47–60）。

【1】次の 3 つの散布図は、東京、O 市、N 市、M 市の 2013 年の 365 日の各日の最高気温のデータをまとめたものである。それぞれ、O 市、N 市、M 市の最高気温を縦軸にとり、東京の最高気温を横軸にとってある。以下のア～オの記述において、正しい場合は○を、正しくない場合は×をつけよ。

（ア） 東京と N 市、東京と M 市の最高気温の間にはそれぞれ正の相関がある。
（イ） 東京と N 市の最高気温の間には正の相関、東京と M 市の最高気温の間には負の相関がある。
（ウ） 東京と N 市の最高気温の間には負の相関、東京と M 市の最高気温の間には正の相関がある。
（エ） 東京と O 市の最高気温の間の相関の方が、東京と N 市の最高気温の間の相関より強い。
（オ） 東京と O 市の最高気温の間の相関の方が、東京と N 市の最高気温の間の相関より弱い。



出典：『過去の気象データ』(気象庁 Web ページ) などにより作成

【2】ある母集団の平均は 5 で、標準偏差は 1 である。この母集団の各要素に 10 を加えたとき、平均と標準偏差はつぎのどれか。最も適するものをひとつ選び記号で答えよ。

(a) 平均 15、標準偏差 1　　(b) 平均 15、標準偏差 5　　(c) 平均 15、標準偏差 11
(d) 平均 10、標準偏差 1　　(e) 平均 10、標準偏差 5

【3】2 つの変量 X; Y について、それぞれ n 個のデータの値が、

$$x_1、x_2、x_3、\cdots、x_n、\quad y_1、y_2、y_3、\cdots、y_n$$

であるとき、X と Y の共分散 s_{xy} は、$\bar{x}$ を X の平均値、$\bar{y}$ を Y の平均値として、

$$s_{xy}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$$

で定義されます（n は正の整数）。

新たな変量 Z のデータ　$z_1、z_2、z_3、\cdots、z_n$ を、a 、b を定数として、次の式

$$z_i = ax_i + b \ (i = 1、2、\cdots、n)$$

で定めるとき、Y と Z の共分散 s_{yz} は、X と Y の共分散 s_{xy} の何倍になるか求めなさい。