

平成 27 年度入試問題における統計“データの分析”の 出題状況と考察

東京理科大学 景山 三平

はじめに

現行の学習指導要領では、小学校 1 年から中学校そして高等学校 1 年までの 10 年間はすべての児童・生徒が系統的に統計教育を受けることになり、統計教育の内容が一層充実している。その中で特に高等学校では統計教育が平成 24 年度から始まり、数学という教科においては唯一の必修科目「数学 I」の中に“データの分析”として記述統計の内容がある。そこではデータを通して分布の数量的理解と視覚的理解の総合化を目指している。すでに 3 年間の経過した。これを受けて平成 27 年度大学入学試験で記述統計の問題を数学 I の範囲として出題することが可能となった。これは従前の数学 B での選択問題としての統計の扱いと些か異なるものである。この意味で平成 27 年度大学入学試験において統計の問題の出題状況やその内容は、今後の高等学校の数学教員の統計教育に対する取り組む姿勢、及び高等学校における統計教育の実施形態に大きな影響を与えると考えている。また選択扱いとなる現在の数学 B の“確率分布と統計的推測”（今春の大学入試センター試験では約 5% の選択率と聞いている）の内容もしかりである。

ちなみに 2015 年 1 月 18 日実施の大学入試センター試験の科目「数学 I・数学 A」では必答の第 3 問に 15 点配点で次のようなものが出題されていた：データのヒストグラム表示に対して（1）第 3 四分位数が属する階級の決定（2）箱ひげ図との対応の理解（3）2 つの箱ひげ図からのデータの変容の理解（4）散布図、平均値、中央値、分散、標準偏差、共分散の値から相関係数の求値。これは現行の学習指導要領の趣旨に合致した問題である。これらの問題についての分析等は、大淵（2015）や深澤他（2015）にも見られる。

本報告は、景山（2015）を改訂したものであり、まず筆者が所属していた広島工業大学で出題された統計内容の 7 種類の問題について解答状況等を分析する。当大学では平成 27 年度入試においては、旧課程の生徒にも配慮することから、統計の問題は入試教科「数学」の中で選択問題の扱いであった。平成 28 年度入試ではこの配慮は不要であるが、統計の問題の出題も含めた扱いが大学内で改めて議論になると思われる。なお、当大学では、数学の問題は 6 年前よりすべて記述式であったが、今回統計の問題（形態 Sa のみ）では選択式になっていた。最後に統計の内容に関する問題の全国的な出題状況に簡単に触れる。

1. 統計の問題とその選択状況

広島工業大学では、数学を科目「数学Ⅱ・数学B」で受験する工学部、「数学Ⅰ・数学A」で受験してもよい学部（情報学部・環境学部・生命学部）がある。その中で今回初めて統計の内容が「数学Ⅰ・数学A」の選択問題として出題されたが、入試で初めての出題であることから受験生による実際の選択状況に関心があった。実際、終わってみると予想以上の選択率であった。この予想以上という認識は高等学校での授業科目「数学Ⅰ」の“データの分析”の授業実施において種々の課題が教育現場で語られていたからである。この内容は大体7時間の授業構成で実施されている学校が多いと聞いているが、この授業時間数では十分ではないと考えている。高等学校の先生方は多分今回の大学入学試験の問題でどのような統計の内容が出題されるかの様子見だと考えられる。

数学が課せられる試験が7回の入試形態で実施された。その7回分の「数学Ⅰ・数学A」の試験問題に出題された統計に関する問題の選択率とその解答状況等を紹介する（実際の問題文は付録を参照）。またもう一方の入試科目「数学Ⅱ・数学B」では統計の問題は出題されていなかったが、確率に関する小問は出題されていた。

(1) 併願推薦入学試験前期日程 Sa (50分) [2014年11/15実施]

- ・大問Ⅰ, Ⅱ番が必答, Ⅲ(確率), Ⅳ(統計)番が選択で, 合計3問を解答。
- ・大問Ⅳ番では, 2変量データに関する9つの散布図に対応する相関係数の値を12個の値の中から選択するもの{選択理由は不問}。
- ・大問Ⅳの選択率: 8.5%

(2) 併願推薦入学試験後期日程 Sb (50分) [2014年12/20実施]

- ・大問Ⅰ, Ⅱ番が必答, Ⅲ(確率), Ⅳ(統計)番が選択で, 合計3問を解答。
- ・大問Ⅳ番では, 3つの箱ひげ図の同時表示をみて7つの設問文に対する正誤の判断及びその判断の理由を自由記述するもの。
- ・大問Ⅳの選択率: 57.1%

(3) 一般入学試験A日程 Aa (80分) [2015年1/31実施]

- ・大問Ⅰ, Ⅱ, Ⅲ番が必答, Ⅳ(確率), Ⅴ(統計)番が選択で, 合計4問を解答。
- ・大問Ⅴ番では, 平均値, 中央値, 最頻値の値をそれぞれ与えた下で, 30個のデータに関する度数分布表の中の2つの(未知の)度数の値を求めるもの。
- ・大問Ⅴの選択率: 96.9%

(4) 一般入学試験A日程 Ab (80分) [2015年2/1実施]

- ・大問Ⅰ, Ⅱ, Ⅲ番が必答, Ⅳ(確率), Ⅴ(統計)番が選択で, 合計4問を解答。
- ・大問Ⅴ番では, 2つのグループのそれぞれの平均値, 標準偏差の値のみを与えて, 全体の平均値と分散の値を求めるもの{具体的なデータの値は与えていない}。
- ・大問Ⅴの選択率: 63.9%

(5) 一般入学試験B日程 Ba (80分) [2015年2/21実施]

- ・大問Ⅰ, Ⅱ, Ⅲ番が必答, Ⅳ(二項定理), Ⅴ(統計)番が選択で, 合計4問を解答。
- ・大問Ⅴ番では, 9つの2次元データの表を与えて, その散布図の作成, それぞれの変

量の平均値，積和，相関係数の値，を求めもの。

- ・大問 V の選択率：34.0%

(6) 一般入学試験 B 日程 Bb (80 分) [2015 年 2/22 実施]

- ・大問 I, II, III 番が必答，IV (確率)，V (統計) 番が選択で，合計 4 問を解答。
- ・大問 V 番では，2 つの未知数を含む 14 個の整数値のデータをある条件下で与え，その未知のデータの値の決定，全体の平均値とデータの範囲，を求めるもの。
- ・大問 V の選択率：43.4%

(7) 一般入学試験 C 日程 (50 分) [2015 年 3/17 実施]

- ・大問 I, II 番が必答，III (確率)，IV (統計) 番が選択で，合計 3 問を解答。
- ・大問 IV 番では，10 個のデータの中で 4 個の値を具体的に与え，それらより小さい 2 つのデータの平均値と標準偏差，それらより大きい 4 つのデータの平均値，さらに全体の標準偏差の値も与えた下で，(1)全体の平均値，(2)中央値，(3)第 1 四分位数，(4)大きい方の 4 つのデータに対する分散の値，を求めるもの。
- ・大問 IV の選択率：94.7%

2. 解答状況

今回の統計の問題は，確率の問題よりは全般的により教科書内容に準拠したものであったので，普通に教科書で学んでいれば受験生も取り組み易かったと思われる。ただ試験問題では当然であろうが，2 つの選択問題から 1 題選択する際にはそれぞれの問題自体の取り組み易さが，統計の問題の選択率に影響を与えていたとも判断できる。

2.1 Sa・Sb 問題について

Sa では統計の問題を選択した者が少なかった。受験者数の 1 割にも満たない。さらに解答状況も悪く最高点でも配点の 3 分の 1 の得点であった。これは相関係数についての数量的理解と視覚的理解の総合化が不十分であることを示している。また相関係数が区間 $[-1, +1]$ の値をとることを理解していない者が殆どであったことに驚くと共に高校現場の統計教育の実態を想像した。一方，Sb は Sa と 1 ヶ月以上経って実施されたので，Sb の受験者は Sa の統計の問題の内容をみて，これは教科書を復習しておけば解けると判断し，適切な試験対策を練った可能性が予想され，それが選択率の顕著な増加に繋がったと考えられる（確率の問題は計算ばかりで楽しくはないから）。Sb の受験者の 3 割が Sa の受験者であったが，残り 7 割は初めての受験者で（Sa 受験者に比べ）成績は極めて良かった。箱ひげ図の問題に対して満点の受験者もいた。ただ解答に際しての判断理由の記述には選択した多くの受験者は戸惑ったようである。これらの記述要求はセンター試験での解答方式との差から生じたものだろう。

2.2 Aa・Ab 問題について

Aa では得点の平均が配点の 7 割近くで満点も何人もいた。教科書的な問題であったので解答し易かったのかも知れない。中央値や最頻値の定義の理解は十分であったが，複数の

度数の組合せまで思考が及ばなかった受験者が 3 分の 1 位いた。これは定義に従って“可能なものをすべて” 探すという数学的な考察ができていないことを意味している。最近の生徒さんは答えが一つ見つかりとすぐ思考停止する人が多いように思う。Ab（ある教科書に同様な問題がある）で全体の平均値を求めることが 8 割ぐらいは出来ていたが、一方全体の分散の値を求める問題は全受験者の中で一人しか出来ていなかったのには驚いた。そのため統計の問題の得点平均が配点の 4 割弱であった。これは計算するだけの問題で、現行の学習指導要領の改訂趣旨にマッチしていないため、高校現場で計算練習があまりなされていなかったのが原因かも知れない。このために敬遠されて選択率も下がったとみている。それとも分散そのものの理解が不十分であるのかも知れない。

2.3 Ba・Bb 問題について

Ba では A 日程入試に比べ統計の選択率がかなり下がった。その理由はもう一つの選択問が二項係数の計算に関するものだったのでその方がやり易いと判断した結果であろう。その間では満点が数人いた。一方統計[V]（選択率 34%）では満点の受験者は居なかった。特に相関係数を求める小問に正解者が全く居なかったのには驚いた。そのため配点の 6 割程度の出来であった。Sa の解答でも感じていたが、どうも相関係数についての理解に課題があるように思う。共分散との関連で的確に学んでほしい。今回の大学入試センター試験での相関係数のことでも現場教師から難しかったという声を聞いている。やはり最小限の数量的理解は醸成してほしいものである。また散布図の作成も 3 分の 1 の選択者は出来ていなかった。他にデータの数値が分数で与えてあったがこれには受験生は戸惑ったと感想が届いている。Bb では、Ba の統計の問題に比べ選択率がすこし上がった（確率の問題が馴染みにくかったかも知れない）が配点の 6 割程度の出来であった。特にデータが相異なるという条件を吟味しないため未知のデータの値をすべて正確に決定することが出来なかった者が多かった。またデータの範囲は中学校の教材であるが 4 分の 1 の選択者しか出来ていなかったのには非常に驚いた。

2.4 C 問題について

C 日程入試問題では統計の問題の選択率がかなり高かった。実際、科目「数学 I・数学 A」での受験者数は多くはなかったが、そんな中で一人の受験生を除いて全員が統計の問題を選択していた。これは広島工業大学の今年の入試では統計の問題が出題されるという評判が広まった結果であろう。この間に対する得点平均は配点の 6 割弱であった。小問(1)～(3)については、特性値の意味も考えながら思考結果を種々記述していたが、小問(4)の分散の値の導出には正解者は皆無であった。これは、高等学校で分散の定義式と実際と同値の簡便な計算式の意味理解が十分になされていなかったことに原因があると思う。このことは丁度 Ab の分散求値の問題の場合と同様な状況を示していた。それらの状況を高校教育現場で再認識してほしいものである。

3. 課題

筆者の日頃の課題意識と今回の広島工業大学における入学試験結果の状況を合わせて統計教育に関する3つの課題を記述する。

(1) 筆者は現在高校教育現場で使用中の教科書「数学 I」において、平成 26 年度版の占有率が 100%の 5 社のテキストの内容“データの分析”にすべて目を通したが、各社様々で工夫もあった。その記述の内容には少し課題があるにしても、生徒が大きな混乱をきたすようなものはないと判断した(景山, 2012a)。そのように考えると今回の入学試験での解答状況から結局、教育現場での先生方の授業方法の中に何らかの課題があるのではと思っている。統計専門家による先生方へのさらなる支援が必要であると認識している(景山, 2012b)。このような状況下で現在、すでに次期の学習指導要領改訂のための審議が始まっているが、学校現場での統計教育は一体どういう方向に進むのでしょうか。教科「算数」「数学」の中では統計教育はさらに充実する予感もありますが、それに伴った現場の先生方の資質向上の施策は大丈夫でしょうか。

(2) 教科書を記述内容に沿って通り一遍に教える(教えさせられている)先生方も多いと感じている。仮にそうだとするとまた現場で生徒へただ覚えて計算することを要求するような古い統計授業スタイルとなる可能性が強まる。これは好ましい状況ではない。記述されている概念や特性値等の必要性和その意味の理解ができないと、結局実際には使えないし活用できないということを十分に認識してほしいと考えている。少なくとも箱ひげ図を利用して分布の把握ができるようにはなりたいものである。

(3) 現行の学習指導要領下での統計教育は、統計の普及からも当然ではあるが、その重点が従前の計算や整理中心から諸特性値の必要性やその意味の理解に移っているので、統計教育に関心・理解のある教員がいない大学では入試問題の作成にあたって統計の問題の前述趣旨に沿った出題は容易ではないと感じた。平成 28 年度個別入試からは統計の問題は必修として出題できる環境であるが、各大学において数学の試験の中でその問題数との関連において果たして実際に統計の問題が出題されるか疑問に思っている。各大学での統計の問題の出題に向けた前向きな検討を期待したい。

4. 他大学での出題状況

今回は新旧両課程で学んだ受験生への配慮の年であったので、統計の問題がそんなに多く出題されるとは考えていなかった。実際、平成 27 年度入試では、出題範囲を旧課程と新課程の共通部分からとした大学が多かった。これらの大学では「データの分析」の内容は当然出題の範囲外となる。数学 II/B の「確率分布と統計的な推測」も選択問題として出題されている大学もありますが、ここでは筆者が知り得た情報で数学 I「データの分析」について主に記述する。そこでの統計の問題の実際の選択率は気になるところある。いずれにしても、来春の大学入学試験が本格的な統計の内容「データの分析」の出題元年になることでしょう。

一橋大学(全学部; 三角関数との選択問題): 3 種類の未知の数を含んだ 10 個の 2 変

量データに対して、それぞれの分散の比、相関係数の値、ある条件下でそれらの3つの数の決定の問題。

同志社大学（政策，スポーツ健康，文化情報学部；選択問題）：2変量の10個の変数データに対して5つの関係式を与えた上で，1次変換後の変量も含め，それらの平均，関係式の証明，相関係数の求値。

慶応義塾大学（商学部，総合政策学部）では，統計に関する文章を読んだり，統計データを見て記述するような論文形式的な要素の濃い問題が出題されたようである。

岡山理科大学では，確率の問題が必答問題として出題されていたようである。

鹿児島大学では，選択問題として確率の問題（新課程：理・医・歯・工・教），標本平均の平均・分散の導出及び信頼区間の問題（旧課程：理・医・歯・工）が出題されていた。

他に確率に関する問題が，岐阜薬科大学，久留米大学，明治大学，同志社大学で出題されているようである。今回，学ぶ科目が変わった期待値の扱いで不適切な大学も見られた。

参考文献

大淵智勝（2015）. 大学入試における統計分野の出題と得点分布とこれからの入試. 統計教育の方法論ワークショップ 統計数理研究所 3月（統計数理研究所共同研究リポートリポート 335, 統計教育実践研究, 第7巻, 85-88）.

景山三平（2012a）. 新学習指導要領に基づく高校教科書「数学 I」の統計記述内容及びその評価. 広島工業大学紀要 教育編, 第11巻, 61-66.

景山三平（2012b）. 「数学 I」の新内容“データの分析”について. じつきょう 数学資料 N.64, 1-3.

景山三平（2015）. ある私立大学における平成27年度入試問題「統計」の分析と課題. 統計教育の方法論ワークショップ 統計数理研究所 3月（統計数理研究所共同研究リポートリポート 335, 統計教育実践研究, 第7巻, 81-84）.

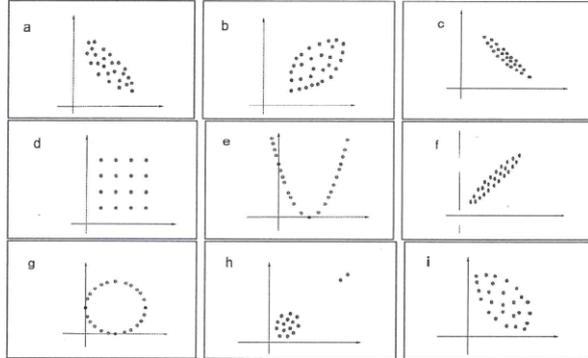
深澤弘美・櫻井尚子・和泉志津江（2015）. 大学入試センター試験と海外の出題傾向. 統計教育の方法論ワークショップ 統計数理研究所 3月（統計数理研究所共同研究リポートリポート 335, 統計教育実践研究, 第7巻, 69-72）.

付録

平成27年度入学試験問題での統計「データの分析」に関して広島工業大学で扱った実際の7つの問題を原文のまま（縮小版）抜粋して以下に転載する（Sa, Sb, Aa, Ab, Ba, Bb, Cの順に）。

[IV] (選択問題)

2つの変量の9種類のデータの散布図 a ~ i に対応する相関係数の値を下の (1) ~ (12) の中から選び、その番号を解答用紙の所定の欄に書け。

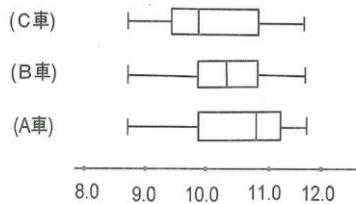


- (1) $r = -1.4$ (2) $r = -0.95$ (3) $r = -0.7$ (4) $r = -0.5$
 (5) $r = -0.2$ (6) $r = 0$ (7) $r = 0.55$ (8) $r = 0.8$
 (9) $r = 0.95$ (10) $r = 1$ (11) $r = 1.5$ (12) $r = 2$

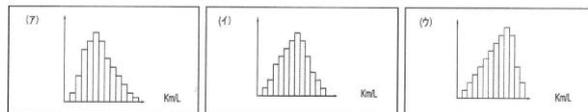
Sa - 数 1-2

[IV] (選択問題)

3台の車の燃費データ (Km/L) についての下の箱ひげ図において、下記の (1) ~ (7) の説明が正しければ ○ を、誤りであれば × を、その理由とともに解答用紙の所定の欄に書け。



- (1) 3台の車のデータの範囲はすべて同じである。
 (2) 3台の車のデータの四分位範囲はすべて同じである。
 (3) A車とC車のデータの中央値の差はB車のデータの四分位範囲と同じである。
 (4) 3台の車のデータの分布はすべて下図の (イ) の形状である。
 (5) A車のデータの分布は下図の (ア) の形状で、C車のデータの分布は下図の (ウ) の形状である。
 (6) A車のデータの分布は下図の (ウ) の形状で、C車のデータの分布は下図の (ア) の形状である。
 (7) B車のデータの分布は下図の (イ) の形状である。



Sb - 数 1-2

[V] (選択問題)

30人のクラスで10点満点のテストを行い、その結果は次の表の通りである。

得点	0	1	2	3	4	5	6	7	8	9	10	計
人数	0	0	2	4	5	a	b	2	3	4	3	30

次の問いに答えよ。

- (1) $a + b$ の値を求めよ。
- (2) 得点の平均値が6点のとき、 (a, b) を求めよ。
- (3) 得点の中央値が5.5点のとき、 (a, b) を求めよ。
- (4) 得点の中央値が6点のとき、 (a, b) を求めよ。
- (5) 得点の最頻値が6点のとき、 (a, b) を求めよ。

Aa - 数1-2

[V] (選択問題)

12個のデータがある。そのうちの6個のデータの平均値は4, 標準偏差は3であり、残りの6個のデータの平均値は8, 標準偏差は5である。

次の問いに答えよ。

- (1) 全体の平均値を求めよ。
- (2) 全体の分散を求めよ。

Ab - 数1-2

[V] (選択問題)

2つの変量 (X, Y) が次の9つのデータ (x_i, y_i) ($i = 1, 2, 3, \dots, 9$) をとるとき、次の問いに答えよ。

i	1	2	3	4	5	6	7	8	9
X	-1	$-\frac{1}{2}$	0	$\frac{1}{2}$	1	$\frac{3}{2}$	2	$\frac{5}{2}$	3
Y	4	$\frac{9}{4}$	1	$\frac{1}{4}$	0	$\frac{1}{4}$	1	$\frac{9}{4}$	4

- (1) このデータの散布図を書け。
- (2) 変量 X, Y のそれぞれの平均値 \bar{x}, \bar{y} の値を求めよ。
- (3) $x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4 + x_5y_5 + x_6y_6 + x_7y_7 + x_8y_8 + x_9y_9$ の値を求めよ。
- (4) 相関係数を求めよ。

Ba - 数1-2

[V] (選択問題)

次のような相異なる14個の正の整数値のデータがある。

24, 19, 4, 21, 2, 8, 26, 5, 7, a , 28, 10, b , 30

ただし、 $9 < a < b$ とする。中央値が14であるとき、次の問いに答えよ。

- (1) すべての組 (a, b) を求めよ。
- (2) 平均値を求めよ。
- (3) 範囲を求めよ。

Bb - 数1-2

[IV] (選択問題)

10 個の卵がある。そのうち 4 個は 55 g, 56 g, 54 g, 55 g で、それらより小さい 2 個の平均が 50 g で標準偏差が 1 g, それらより大きい 4 個の平均が 60 g である。また、全体の標準偏差は 4 g である。

次の問いに答えよ。

- (1) 全体の平均値を求めよ。
- (2) 10 個の卵の重さの中央値を求めよ。
- (3) 第 1 四分位数を求めよ。
- (4) 大きい方の 4 個の卵の重さの分散を求めよ。