高等学校新学習指導要領における統計の扱いについて

武　沢　護 (takezawa mamoru)
早稲田大学高等学院/大学院教職研究科*

1　次世代の数学教育の方向性について

次世代の教育の方向性について，アメリカ (USA) では"The Partnership for 21st Century Skills" というプロジェクトのもと，キーワードとして批判的思考力（Critical thinking），問題解決能力（Problem solving），コミュニケーション能力（Communication and Collaboration）の育成が重点化されている．

一方，ヨーロッパ (EU) では PISA（OECD）での調査でも知られるように読解力（Reading），数理的能力（Mathematical and Science literacy），問題解決能力（Problem solving）の育成が問われている．

わが国においても新学習指導要領のもと，次世代の教育の改善が図られ，今回の高等学校数学での数学における「データの分析」の必修化を始めとした統計的な数学の扱いの重点化はこの流れの一環とみていいだろう．この視点は，早稲田大学が 1913 年以来教旨として学問の独立（to uphold the independence of learning），学問の活用（to promote the practical utilization of knowledge），模範国民の造就（to promote good citizenship）を掲げてきたものと通底し，この教えは 21 世紀に向けて輝きを失っていない．

1.1　新しい学習指導要領における扱いと問題点

中学校において，新しい学習指導要領（中学の全面実施は 2012 年から．移行措置として 2009 年度から）では，領域構成において現行の「数と式」「図形」「数量関係」の 3 領域から確率・統計に関する領域「資料の活用」を新設するとともに,「数と式」「図形」「関数」「資料の活用」の 4 領域に再構成した．第 1 学年ではヒストグラムや代表値を用いて資料の傾向を読み取る．また，第 2 学年では確率を用いて不確定な事象をとらえ説明することを学習する．さらに第 3 学年では標本の傾向を調べることで，母集団の傾向を読み取ることを簡単な標本調査を行うことで理解できるようにするとなっている．

一方，高等学校においては次のようになった．ただし ( ) 内は単位数である．

◎現　行：数学　(3)，数学　(4)，数学　(3)，数学 A(2)，数学 B(2)，数学 C(2)，数学基礎 (2)
◎新課程：数学　(3)，数学　(4)，数学　(5)，数学 A(2)，数学 B(2)，数学活用 (2)

特に，確率・統計分野に関する具体的な内容は次の通りである．

・数学　(必履修)：データの分析（データの散らばり，データの相関），課題学習
・数学 A：場合の数と確率（選択），課題学習
・数学 B：確率分布と統計的な推測（選択）
・数学活用：社会生活における数理的な考察

※数学 B の内容：確率分布（確率変数，二項分布），正規分布，統計的な推測，母集団と標本，統計的な推測の考え方（信頼区間）

*takezawa@waseda.jp

なお，他教科においては理科ではもちろんのこと，地理，公民（政治経済）などにおいても「資料の収集，処理や発表などにあたっては，コンピュータや情報通信ネットワークなどを積極的に活用…」とある点にも注目したい．

1.2　問題点（教科書と大学入学試験）

前に述べたように，21 世紀を担う子どもたちがさまざまなデータに触れながら批判的な精神をもった市民として生きていくためにもこの分野は重要であると考えるが，新しい指導要領のもと高等学校の授業や大学入試で十分に扱われるか非常に疑問である．ここでは問題点をいくつか指摘しよう．

まず，取り上げられるデータの質とそれに関連するコンピュータの扱いである．必修である数学　の教科書（数研出版）[?] での扱いから感じることは，コンピュータの利用を前提としていないためデータ量が少なく興味深いデータサンプルを扱っていない．この分野はそもそも応用数学的側面が強いのだから，扱うデータの面白さが非常に重要になってくる．またコンピュータでの扱いは章の最後に表計算ソフトの扱いはあるがあくまでも副次的な扱いである．

次に，数学 B での扱いである．数学 B では「確率分布と統計的な推測」,「数列」,「ベクトル」の３つの単元から２つの単元を選択することなっているが，各高等学校での選択パターンによってはこの統計的な数学は選択されない可能性が高い．そして，東京大学を始めとするいくつかの有力大学は数学 B の確率分布の単元を入学試験から除くと早々と発表しているのも問題である．

これら抱える問題点を解決する方策として，次の視点から検討していくことが重要であると考える．

・諸外国のカリキュラムならびに教科書の構成の調査
・教育的に興味深いデータサンプルの収集やそれらを活用した教材開発
・大学入試にふさわしい問題群の開発（数　や融合問題，過去の問題の研究）
・コンピュータの活用（Excel，R，Mathemaica 等）

2　英国の統計に関するカリキュラム

イギリス（主にイングランド）では，1988 年の教育改革によりナショナルカリキュラムが制定され，その到達度を評価するための全国一斉テスト（SAT）が実施されている． そして，セカンダリスクール最終学年（16 歳、わが国では中学卒業程度）は全員 GCSE(General Certificate of Secondary Education:一般中等教育資格試験) を受験する．これには必修科目と選択科目がある． さらに大学進学を希望する者はシックスフォーム（Sixth form）に 2 年在籍し，その後，GCE − AS/A2（General Certificate of Education − Advanced Subsidiary Level/Advanced Level）試験を受ける．

2.1　GCSE 以降

GCSE 以降に関して, 必修数学（Core Mathematics) の他に，発展数学（Further Pure Mathematics)，統計（Statistics)，力学（Mechanics)，意思決定数学（Decision Mathematics) などがある．特に，統計のカリキュラムは S1，S2，S3，S4 と 4 つのモジュールがあり，それに準拠した教科書が作成されている [?].

S1：Representation of data, Measures of location, Measures of spread, Probability, Permutations and combinations, Probability distributions, The binomial and geometric distributions, Expectation and variance of a random variable, Correlation, Regression.

S2：Continuous random variables, The normal distribution, The Poisson distribution, Sampling, Hypoth-

esis testing, Errors in hypothesis testing.
S3：Continuous random variables, Linear combinations of random variables, Confidence intervals and the t-distribution, Testing for differences between populations, Chi-squared tests.
S4：Probability, Non-parametric tests, Probability generating functions, Moment generating functions, Estimators, Discrete bivariate distributions.

2.2 英国 S1 レベルからの出題

次の問題はイギリスの S1 レベルの教科書 [?] に掲載されている問題を参考にして作成したものである．このレベルはわが国の新学習指導要領での数学　レベルに近い．

1. 次の表はある店の 12 週の売上高を表したものである．

週	1	2	3	4	5	6	7	8	9	10	11	12
売上 x_i 万円	5.5	4.2	5.8	9.1	3.8	4.6	6.4	6.2	4.9	5.9	6.0	4.1

ただし，和 $\sum_{i=1}^{12} x_i = 66.5$, 平方和 $\sum_{i=1}^{12} {x_i}^2 = 390.97$
(1) 平均と分散を求めよ．
(2) メジアンと第 1 四分位数，第 3 四分位数を求めよ．
別の店の同時期の売上高を調査した結果, 次のデータが得られた．
平均：5.48，分散：0.726，メジアン：5.6，四分位範囲：1.5
(3) 二つの店の平均と分散を比較せよ．
(4) 扱うデータによって，平均よりもメジアンの方が代表値としてふさわしい場合がある．この理由を述べよ．

2. ある工場では 100m のケーブルを生産している. あるとき品質管理部門が任意の 20 本を検査したところ次のような結果になった．
和 $\sum_{i=1}^{20} x_i = -86$, 平方和 $\sum_{i=1}^{20} {x_i}^2 = 4281$
ただし, 単位 99.84m の場合は変量 x を-16 とする．
(1) 変量 x の平均 $\overline{x}$ と標準偏差 σ を求めよ．
(2) この 20 本のケーブルの平均と標準偏差を求めよ．
(3) このデータの中に-47 というデータが紛れていた．この管理部門では $|x - \overline{x}| > 2\sigma$ のデータはエラーとみなすことにする．-47 がエラーであることを確認せよ．
(4) この中に-47 という誤りのデータが紛れていた. この-47 を除いた変量 x の平均および標準偏差を求めよ．

3. ある会社では 20 人の従業員が働いている. この従業員の時給は次の通りである．
820, 340, 650, 520, 330, 750, 790, 550, 420, 890, 560, 450, 400, 460, 1200, 620, 700, 780, 350, 420
(1)　和 $\sum_{i=1}^{20} x_i = 12000$, 平方和 $\sum_{i=1}^{20} {x_i}^2 = 8156800$
があたえられるとき, この従業員の時給の平均と標準偏差を求めよ．

(2) この従業員は、雇用者から次の２つの提案を受けた.

ア　一律時給 60 円の増額　　　イ　一律 10 %の増額

この二つの提案の時給の平均および標準偏差を求めよ.

(3) 雇用者従業員の立場でこの２つの提案について簡潔に述べよ.

4. 20 人の新生児の身長と頭囲のデータが次のように得られた。
   頭囲 x：31,32,33.5,33.5,34,34,34,34,34,34.5,34.5,35,35,35,35,35.5,36,36.5,36.5,37.5
   身長 y：45,49,49,51,47,50,51,52,53,51,52,50,51,52,54,52,53,52,53,51
   ただし、統計データは次の通りである.
   $\sum_{i=1}^{20} x_i = 691$, $\sum_{i=1}^{20} {x_i}^2 = 23917$, $\sum_{i=1}^{20} y_i = 1018$, $\sum_{i=1}^{20} {y_i}^2 = 51904$, $\sum_{i=1}^{20} x_i y_i = 35212.5$
   (1) この変量 x と変量 y との相関係数を求めよ.
   (2) このデータのうち，x=34, y=51 を間違って x=51, y=34 と記録してしまった．統計データや相関係数はどのように変わるか.

3　統計と他分野と関連させた問題

ここでは，数学　分野だけでなく他の分野（数学　、数学 B，数学　）と関連・融合させた問題を取り上げ, 高等学校の授業や大学入試で扱うことが可能な問題について検討する.

3.1　数学

次の問題 1 は 2 次方程式と相関係数を融合させたものである. 問題 2 は変数変換．問題 3，4 はチェビシェフの不等式に関連している.

1. 不等式 $\sum_{i=1}^{n}\{t(x_i-\overline{x})-(y_i-\overline{y})\}^2 \geq 0$ を利用して，相関係数 r が $-1 \leq r \leq 1$ となることを証明せよ.
   ただし t は任意の実数．$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$,　$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ である.

2. ２組の標本 $x_1, x_2, \ldots, x_n$,　$y_1, y_2, \ldots, y_n$ があって, その間に $y_i = ax_i + b$ (a, b は定数，$i = 1, 2, 3, \ldots, n$ の関係がある．おのおのの標本平均を $\overline{x}$,$\overline{y}$, 標準偏差を S_x, S_y とするとき, $\overline{x}$ と $\overline{y}$ の関係および S_x と S_y との関係を求めよ.（愛知教育大）

3. 変量 x の資料の総数を N, x のとる値 $x_1, x_2, \ldots, x_n$ の平均を $\overline{x}$, 標準偏差を σ とすれば $|x-\overline{x}| < k\sigma (k > 0)$ の範囲に含まれる資料の個数は $N(1-\frac{1}{k^2})$ 以上であることを証明せよ.（神戸大）

4. n 個のデータ $x_1, x_2, \ldots, x_n$ の相加平均を $\overline{x}$，標準偏差を σ とするとき
   (1) 平均からの偏差 $x-\overline{x}$ の平方の総和と任意の数 c からの偏差 $x-\overline{x}$ の平方の総和の大小を比較せよ.
   (2) $|x-\overline{x}| \geq k\sigma$ をみたす x_i の個数を m とすると,$m \leq \frac{n}{k^2}$ となることを示せ. ただし，$k > 0$,$\sigma > 0$ とする.

3.2　数学 B

次の問題は相関係数の問題をベクトルと関連させた.

1. ベクトル $\overrightarrow{X} = (X_1, X_2, \ldots, X_n)$, $\overrightarrow{Y} = (Y_1, Y_2, \ldots, Y_n)$, に対して次の問いに答えよ.
   (1) $(\overrightarrow{X} \cdot \overrightarrow{Y})^2 \leq |\overrightarrow{X}|^2 |\overrightarrow{Y}|^2$ を証明せよ.
   (2) データ $d_x = (x_1, x_2, \ldots, x_n)$，$d_y = (y_1, y_2, \ldots, y_n)$ に対して，$\overline{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i$，　$\overline{y} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} y_i$ とする．このときのデータ d_x, d_y の相関係数 r が $-1 \leq r \leq 1$ となることを証明せよ.

4　過去の入試問題から

ここでは過去の大学入試で出題された問題をいくつか列挙し今後の参考とする．これらのほとんどは新しいカリキュラムでの数学 B の確率分布や数学　の問題に関連している.

1. ある容器の中に，1 から n までの数字を一つずつ書いたボールが n 個入っている．この容器からボールを 1 個ずつ 2 回取り出したときの最大の数を X とする．ただし，最初に取り出したボールは容器の中に返すものとする.
   (1) $X \leq k (k = 1, 2, \cdots, n)$ となる確率を求めよ.
   (2) X の確率分布および X の平均値を求めよ.
   (3) X の分散 V(x) とするとき，$\displaystyle\lim_{n\to\infty} \frac{V(x)}{n^2}$ を求めよ.（79 島根医大）

2. ある人が公平なコインを投げて，表が出れば x 軸上を 1 進み，裏が出れば 1 後退するものとする．原点を出発点として，n 回コインを投げた後にこの人がいる位置を確率変数 X_n とし，$X_n = k$ である確率を $P_n(k)$，X_n の分散を V_n とする.
   (1) $P_n(k)$ を $P_{n-1}(k-1)$, $P_{n-1}(k+1)$ で表せ.
   (2) V_1，V_2，V_3 を求めよ.
   (3) V_n の値を類推し，数学的帰納法で証明せよ.（80 三重大医）

3. 確率変数 X の確率分布が次の表のように与えられている.

X	x_1	x_2	$\cdots$	x_n	計
確率	p_1	p_2	$\cdots$	p_n	1

   また，X の期待値を m，分散を σ^2 とすれば，次の不等式が成立することを示せ.（$\sigma \geq 0$)
   確率　$P(|X - m| > k\sigma) \leq \dfrac{1}{k^2}$
   （ただし，k は任意の正の定数）　　（80 武蔵工大）

4. 区間 $\lfloor -\dfrac{m}{2}, \dfrac{n}{2} \rfloor$ のすべての値をとる確率変数 X の確率密度関数が

$$f(x) = \begin{cases} \dfrac{1}{m}x + \dfrac{1}{2} & -\dfrac{m}{2} \leq x \leq 0 \text{ のとき} \\ -\dfrac{1}{n}x + \dfrac{1}{2} & 0 \leq x \leq \dfrac{n}{2} \text{ のとき} \end{cases} \tag{1}$$

により与えられている．X の平均値が $\dfrac{2}{3}$ のとき，次の問いに答えよ.
(1) m，n の値を求めよ.
(2) X の分散を求めよ.
(3) t の 2 次方程式
$t^2-4t+2X+1=0$ の 2 つの解がともに正である確率を求めよ.（81 静岡大理, 工）

5. ある県の 20 歳の男子を母集団とするとき，その身長は平均 165　，標準偏差 4　の正規分布をするという．いま，

F(x)= $\displaystyle\int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$

とするとき，x と F(x) の対応が下のように与えられているものとする.

x	0.126	1.000	1.645	2.000
F(x)	0.050	0.341	0.450	0.477

(1) この母集団の身長の上位 5 %の者は何　以上と考えられるか.
(2) この母集団から無作為に 64 人の標本を抽出したとき，その標本平均が 164　以上かつ 166　以下である確率を求めよ.（79 宮崎大理系）

参考文献

[1] 検定教科書「数学　」，数研出版，2011.

[2] Statistics, ADVANCED MATHEMATICS，Cambridge University Press UK，2010.

[3] Statistics, MEI STRUCTURED MATHEMATICS，HODDER EDUCATION UK，2004.