茨城大学教育実践研究 31(2012), 1-15

統計判断における学習者の誤りに関する問題

小　口　祐　一*

(2012 年 9 月 15 日受理)

A Survey of Systematic Errors in Statistical Judgment

Yuichi OGUCHI

キーワード：統計判断，誤概念，バイアス，ヒューリスティック

本研究の目的は，「近年のわが国の大学生にみられる統計判断の誤りには，どのようなものがあるか。」という問いに答えることであった。まず，先行研究をレビューし，「結果の事前確率の無視」(Insensitivity to prior probability of outcomes)，「標本の大きさの無視」(Insensitivity to sample size)，「偶然の誤概念」(Misconception of chance)，「予測可能性の無視」(Insensitivity to predictability)，「妥当性の幻想」(The illusion of validity)，「回帰の誤概念」(Misconception of regression)など，統計判断における誤りについて概観した。次に，私立大学文学部の学生 68 名を対象に，統計判断に関する調査を実施した。その結果，近年のわが国の大学生にみられる統計判断の誤りには，「分布の形状の誤判断」，「無作為化の対象の誤判断」，「頻度の誤判断」，および先行研究でみられた「標本の大きさの無視」が存在することを確認した。これらの誤判断は，学習者が，適切判断に必要な統計の知識を適用せずに，「代表性ヒューリスティック」を発動してしまうことが，一つの要因となっていると考えられる。これらの誤判断を修正するための教授方略を構築することが，今後の課題として残された。

統計判断における誤概念とバイアス

統計の知識を適用した適切判断がなされない背景として，学習者が狭く偏った範囲の経験から誤った知識を獲得し，そのような誤った知識を優先することにより，正しい知識の適用が阻害されてしまうことが挙げられる。これらは，「誤概念」や「ル・バー」という用語として概念化されている。誤概念とは，「現代科学における正しい概念とは，必ずしも一致しない概念」のことであり，ル・バーとは，「現代科学と必ずしも一致しないという意味で誤りであるが，ある程度の一般性を持ったルール」のことである（麻柄ほか，2006)。統計判断においては，「偶然の誤概念」や「回帰の誤概念」など，学習者が保持している誤概念の影響が指摘されている。ここでは，先行研究で指摘された誤概念，および学習者が保持する偏った判断の傾向（以降，「バイアス」という。）について概観する。

---

*茨城大学教育学部

統計判断に関する研究において，一般にバイアスは，誤った判断をもたらす傾向として用いられる。

Kahneman *et al.*（1982）は，統計判断に影響を与える誤概念，およびバイアスについて，「結果の事前確率の無視」（Insensitivity to prior probability of outcomes），「標本の大きさの無視」（Insensitivity to sample size），「偶然の誤概念」（Misconception of chance），「予測可能性の無視」（Insensitivity to predictability），「妥当性の幻想」（The illusion of validity），「回帰の誤概念」（Misconception of regression）などの存在を指摘している。

タクシーが，夜間のひき逃げ事故に関与していた。市内では，2 つのタクシー会社，緑タクシーと青タクシーが営業している。あなたには，次のデータが与えられる。

(a) 市内のタクシーの 85%は緑タクシーであり，15%が青タクシーである。

(b) 目撃者は，タクシーが青色であったと証言した。裁判所は，夜間に事故が起きた同様の状況における目撃者の証言の信頼度を検討し，目撃者が 2 色を正確に見分ける確率は 80%であり，色を見間違う確率は 20%であったと結論を下した。

事故に関与したタクシーが青色であった確率はいくつか。

ア．80%

イ．41%

ウ．15%

FIGURE 1 タクシー問題

ある町に 2 つの病院がある。大きい病院では毎日約 45 人の赤ちゃんが生まれ，小さい病院では毎日約 15 人の赤ちゃんが生まれる。ご存知のように，赤ちゃんの約 50%は男の子である。しかし，男の赤ちゃんの正確な割合は日によって異なる。あるときは 50%以上かもしれないし，あるときは 50%以下かもしれない。

1 年間を通して，それぞれの病院は生まれた赤ちゃんの 60%以上が男の子であった日を記録した。あなたは，どちらの病院の方が，そのような日が多く記録されたと考えるか。

ア． 大きい方の病院。

イ． 小さい方の病院。

ウ． どちらもほぼ同じ。

FIGURE 2 産科病院問題

「結果の事前確率の無視」（Insensitivity to prior probability of outcomes）とは，ある事象が起きた結果と事後確率が情報として与えられたとき，条件付き確率，あるいはベイズの定理に基づいて確率を判断しなければならない状況であっても，事後確率だけに基づいて判断してしまう傾向のことで

ある。たとえば，タクシー問題（FIGURE 1）は，タクシーによる事故が起き，事故に関与したタクシーが青色であったという目撃者の証言が結果として与えられている。この状況で，多くの学習者に，「(b) 目撃者の証言の信頼度」の情報だけに基づいて，誤って判断してしまう傾向がみられた（Tversky and Kahneman，1974）。

「標本の大きさの無視」(Insensitivity to sample size）とは，「標本の大きさが大きくなるに従って，標本比率は母比率に近づく。」というルール（以降，「大数の法則」という。）に基づいて可能性を判断しなければならない状況であっても，標本の大きさを考慮せずに，母比率だけに基づいて判断してしまう傾向のことである。たとえば，産科病院問題（FIGURE 2）は，2 つの産科病院で標本の大きさが異なっている。この状況で，多くの学習者に，母比率だけに基づいて，誤って判断してしまう傾向がみられた（Kahneman and Tversky，1972）。

ある市において，6 人の子どもがいるすべての世帯の調査が行われた。72 世帯において，男の子と女の子の誕生日の正確な順序は GBGBBG（女男女男男女）であった。あなたは，調査が行われた世帯のうち，誕生日の正確な順序が BGBBBB（男女男男男男）であった世帯数は何世帯であると推定するか。

ア．72 世帯よりかなり少ない。

イ．72 世帯に近い。

ウ．72 世帯よりかなり多い。

FIGURE 3　6 人兄弟問題

ある大学 1 年生に関して，カウンセラーによって記述された知性，自信，よく本を読むこと，一生懸命勉強すること，好奇心の強さなどの印象が伝えられた。この記述にかかわって，次の 2 つの質問に答えなさい。

(a)　評価 : カウンセラーの記述は，この学生の学力にかかわる印象をどのくらい伝えていますか？大学 1 年生のうちで，何%がこの学生よりよい印象が記述されていると思いますか？

(b)　予測 : この学生が得るだろう GPA（成績評価平均）は，いくつと推定しますか？大学 1 年生のうちで，何%がこの学生よりよい成績を得るだろうと思いますか？

FIGURE 4　GPA 問題

「偶然の誤概念」(Misconception of chance）とは，大数の法則に基づいて可能性を判断しなければならない状況であっても，母集団から無作為抽出された標本は，母集団と類似した性質を持つと判断してしまう傾向のことである。ルール命題で表現するならば，「母集団から無作為抽出された標本は，標本の大きさが小さくても，すべての基本的な性質が母集団と類似している。」（Tversky and

Karneman，1971）（以降，「小数の法則」という。）という誤概念に基づいて判断してしまう傾向のことである。たとえば，6 人兄弟問題（FIGURE 3）は，2 つの例で男の子と女の子の誕生日の正確な順序が異なっている。この状況で，多くの学習者に，母集団の基本的な性質（男女比は 1 : 1 である。）だけに基づいて，誤って判断してしまう傾向がみられた（Kahneman and Tversky，1972）。

「予測可能性の無視」（Insensitivity to predictability）とは，測定値は目的変数の平均値に近づくという性質に基づいて将来を予測しなければならない状況であっても，情報が与えられた他の変数との類似性が高いときには，2 つの変数に相関がみられるはずであると判断してしまう傾向のことである。たとえば，GPA 問題（FIGURE 4）は，ある生徒の学力に関する印象の評価が高ければ，この学生が将来得るだろう GPA の予測も高くなると判断してしまう傾向のことである。この状況で，多くの学習者に，印象の評価と GPA という 2 つの変数の間に相関関係があると解釈して，誤って判断してしまう傾向がみられた（Kahneman and Tversky，1973）。

---

2 組のペアの傾向テストに基づいて GPA（成績評価平均）を予測することが求められた。対象者は，1 組のペアの傾向テスト（創造的思考力と表象的能力）には強い相関があり，もう 1 組のペアの傾向テスト（心的順応性と系統的推論能力）には相関がなかったと告げられた。すべてのテストは，大学での目標達成能力を予測するために役立つと告げられた。どちらのテストの方が，大学での目標達成能力を予測するためにより役立つだろうか。

ア．創造的思考力と表象的能力のテスト。

イ．心的順応性と系統的推論能力のテスト。

ウ．どちらも同じ。

FIGURE 5　相関の強さ問題

飛行訓練で，教官は，訓練生が上手に着陸できて褒めたとき，ほとんどの場合次の飛行では上手に着陸できなかったことに気がついた。一方，訓練生が上手に着陸できなかったので厳しく叱ったとき，ほとんどの場合次の飛行では上手に着陸できたことに気がついた。この教官は，言葉で褒めることは学習に有害であり，言葉で厳しく叱ることが学習に有益であると結論づけている。この教官の結論が妥当であるか，あるいは妥当ではないかを示しなさい。

ア．妥当である。

イ．妥当ではない。

ウ．どちらということはできない。

FIGURE 6　飛行訓練問題

「妥当性の幻想」（The illusion of validity）とは，類似した説明変数に基づいた将来の予測の方が，多様な説明変数に基づいた将来の予測よりも妥当性が高いと誤って判断してしまう傾向のことである。たとえば，相関の強さ問題（FIGURE 5）は，創造的思考力と表象的能力という相関が強い 2 つの説明変数に基づく GPA（成績評価平均）の予測の方が，心的順応性と系統的推論能力という相関がない 2 つの説明変数に基づく予測より妥当性が高いと判断してしまう傾向のことである。この状況で，多くの学習者に，相関が強い 2 つの説明変数に基づく目的変数の予測の方が，相関がない 2 つの説明変数に基づく予測より妥当性が高いと解釈して，誤って判断してしまう傾向がみられた（Kahneman and Tversky，1973）。

「回帰の誤概念」（Misconception of regression）とは，測定値は目的変数の平均値に近づくという性質に基づいて将来を予測しなければならない状況であっても，回帰直線上の理想値より測定値が大きかったときには，測定する直前に遂行した行為によってよい結果が得られたと判断してしまう傾向のことである。ルール命題で表現するならば，「ある行為の直後によい結果が得られたならば，その行為はよい結果が得られる原因である。」という誤概念に基づいて判断してしまう傾向のことである。たとえば，飛行訓練問題（FIGURE 6）は，褒めたり叱ったりする行為と，次回の飛行訓練で成功したり失敗したりした結果との間に相関があったことが述べられている。この状況で，多くの学習者に，直前の行為と結果という 2 つの変数の関係は因果関係であると解釈して，誤って判断してしまう傾向がみられた（Kahneman and Tversky，1973）。

それでは，学校教育で統計の正しい概念を学習すれば，上述した誤概念やバイアスは修正され，学習者は適切判断をするようになるだろうか。Garfield and delMas（2005）は，統計の入門コースを受講した大学生の統計的思考力を評価し，改善するための評価テストを提供している。一つは ARTIST（Assessment Resource Tools for Improving Statistical Thinking）と呼ばれる評価テストであり，11 個のトピックごとに 7〜15 問の選択式問題を提示している。もう一つは，CAOS（Comprehensive Assessment of Outcomes in a first Statistics Course）と呼ばれる総合テストであり，40 分程度で回答できる 40 問の選択式問題を提示している。これらの評価テストにおいて，統計的思考力とは，「統計的探究をすすめる目的と方法を理解している能力」のことであり，統計的探究プロセス（問題設定，データ収集，分析手法の選択，仮説検定などのプロセス）を理解していること，統計的推論をいつ，どのように行ったらよいかを認識していること，問題場面の文脈を理解していることなどを含んでいる。2005 年の秋から 2006 年の春にかけて，アメリカ合衆国の大学生 1470 名を対象に CAOS が実施され，正答率や平均値など基本統計量の情報が提供されている。

たとえば，キャンディー工場問題（FIGURE 7）は，標本の大きさが異なる大きなサイズのバッグと小さなサイズのバッグから，標本比率の散らばりが大きいバッグの方を選択することが求められた。この状況で，多くの学習者に，標本の大きさを考慮せずに，母比率だけに基づいて，誤って判断してしまう傾向がみられた（Garfield and delMas，2005）。

キャンディー工場問題の正答率（正答はウ）は，33.0％であり，最も多かった誤答はオで，その反応率は，49.0％であった。この結果から，統計の正しい概念を学習した後でも，多くの学習者は，「標本の大きさの無視」というバイアスを修正せずに，誤判断すると考えられる。

また，標準偏差の大小判断問題（FIGURE 8）は，分布の形状が異なる 5 つのヒストグラムから，標準偏差が最も小さいものを選択すること（問 14）と，最も大きいものを選択すること（問 15）が

求められた。この状況で，多くの学習者に，ヒストグラムを誤読して，頻度の大小に基づいて，誤って判断してしまう傾向がみられた（Garfield and delMas，2005）。

標準偏差の大小判断問題について，問 14 の正答率（正答はア）は，52.8%であり，最も多かった誤答はウで，その反応率は，25.0%であった。問 15 の正答率（正答はイ）は，50.7%であり，最も多かった誤答はウで，その反応率は，29.8%であった。この結果から，統計の正しい概念を学習した後でも，多くの学習者は，ヒストグラムを誤読し，標準偏差の大小判断を誤ると考えられる。

これらのことから，上述した多くの誤概念やバイアスは，学校教育で統計の正しい概念を学習した後でも容易には修正されず，多くの学習者に，継続して統計判断の誤りがみられると推測される。次項以降では，先行研究において誤判断を修正させる様々な教授方略が検討されてきた「標本の大きさの無視」というバイアスと，グラフを読み取るという統計の基本的な能力にかかわる「標準偏差の大小判断」に焦点をあて，数学的な視点から適切判断に必要な数学の概念を明らかにするとともに，心理学的な視点から誤判断の要因に関する検討をすすめる。

あるキャンディー工場では，茶色のキャンディーを 50%，他の色のキャンディーを 50%生産している。サムはキャンディーがたくさん入っている大きなサイズのバッグを買おうと計画している。ケリーはキャンディーが少しだけ入っている小さなサイズのバッグを買おうと計画している。茶色のキャンディーが 70%以上入っている可能性は，どちらのバッグの方が高いでしょうか。

ア．　サム。なぜなら，大きなサイズのバックには，より多くの茶色のキャンディーを入れることができるから。

イ．　サム。なぜなら，大きなサイズのバックに入っている茶色のキャンディーの比率は散らばりが大きいから。

ウ．　ケリー。なぜなら，小さなサイズのバックに入っている茶色のキャンディーの比率は散らばりが大きいから。

エ．　ケリー。なぜなら，小さなサイズのバッグには，大抵 50%以上の茶色のキャンディーが入るだろうから。

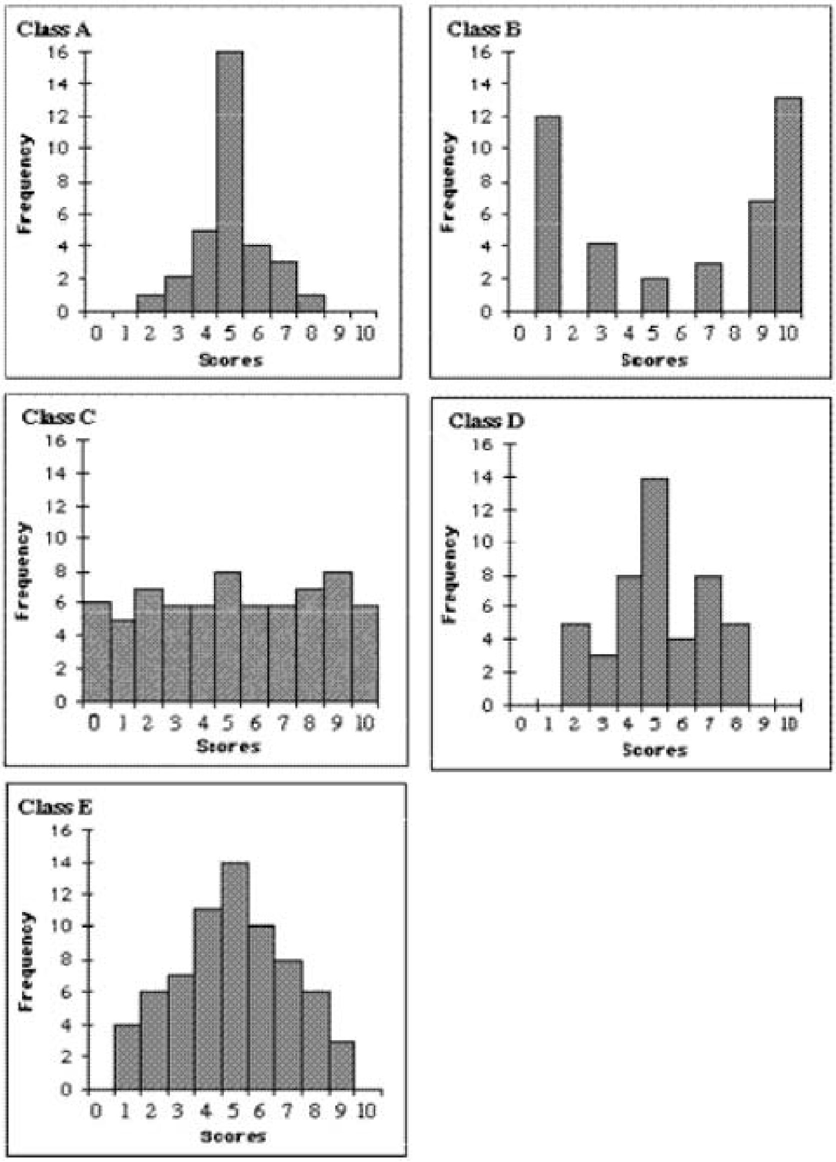
オ．　どちらにも同じ可能性。なぜなら，それらはどちらも無作為標本だから。

FIGURE 7　キャンディー工場問題

大数の法則の無視に関する数学的検討

大数の法則は，無作為抽出という条件のもとで，「標本の大きさが大きくなるに従って，標本比率は母比率に近づく。」というルールである。たとえば，産科病院問題に正答するためには，「標本の大きさが小さければ，標本比率の散らばりは大きい。」という大数の法則の裏命題を適用する必要がある。大数の法則とその裏命題は，次に述べる標本比率の性質を利用して，数学的に説明することができる。

次の 5 つのヒストグラムは, 統計の講義を受講している 5 つのクラスに対して 10 点満点で実施したテストについて，各クラスの得点分布を表示したものである。



問 14.　どのクラスが，標準偏差が最も小さいと推測しますか。

問 15.　どのクラスが，標準偏差が最も大きいと推測しますか。

FIGURE 8　標準偏差の大小判断問題

標本比率の性質

「母比率P，母分散$P(1-P)$の母集団から大きさ n の標本を無作為抽出するならば，標本比率の期待値はP，分散は $\dfrac{P(1-P)}{n}$ になる。」

（解説）

事象 A の母比率が P の母集団において，事象 A の要素を 1，事象 A でない要素を 0 で表す確率変数を X とすると，この母集団の母比率と母分散は,

$$E(X)=0\times(1-P)+1\times P=P$$

$$\sigma(X)=(0-P)^2(1-P)+(1-P)^2P=P-P^2=P(1-P)$$

この母集団から，大きさ n の標本$(x_1,x_2,\cdots,x_n)$を無作為抽出するとき，事象 A の標本比率 R は,

$$R=\frac{x_1+x_2+\cdots+x_n}{n}$$

よって，標本比率 R は標本平均$\overline{X}$と同じになる。大きさ n の標本$(x_1,x_2,\cdots,x_n)$は，それぞれが母集団分布に従う互いに独立な確率変数となるから,

$$E(x_1)=E(x_2)=\cdots=E(x_n)=p \quad \sigma(x_1)=\sigma(x_2)=\cdots=\sigma(x_n)=\sqrt{p(1-p)}$$

したがって,

$$E(R)=E(\frac{x_1+x_2+\cdots+x_n}{n})$$
$$=\frac{1}{n}\{E(x_1)+E(x_2)+\cdots+E(x_n)\}$$
$$=\frac{1}{n}\times np=p$$

$$V(R)=V(\frac{x_1+x_2+\cdots+x_n}{n})$$
$$=\frac{1}{n^2}\{V(x_1)+V(x_2)+\cdots+V(x_n)\}$$
$$=\frac{1}{n^2}\times np(1-p)=\frac{p(1-p)}{n}$$

たとえば, 産科病院問題は, この標本比率の性質を利用して, 次のように説明することができる。

大きい病院で生まれる男の子の赤ちゃんの標本比率 R_1 について，その期待値は $E(R_1)=0.5$，分散は $V(R_1)=0.5\times0.5\div45\fallingdotseq0.0056$ である。同様に，小さい病院で生まれる男の子の赤ちゃんの標本比率 R_2 について，その期待値は $E(R_2)=0.5$，分散は $V(R_2)=0.5\times0.5\div15\fallingdotseq0.0167$ である。したがって，男の子の赤ちゃんが生まれる比率の散らばりは，小さい病院の方が大きいことになる。

大数の法則の無視に関する心理学的検討

上述した統計判断における誤概念やバイアスについて，Kahneman *et al.*（1982）は，人間が用いる方略の一つである「代表性ヒューリスティック」（ある集団の典型的な性質との類似度に基づく判断を促す方略）によって説明している。わが国でも，彼らが指摘する代表性ヒューリスティックによる統計判断への影響が検討されてきた（市川，1996）。

産科病院問題（FIGURE 2）に関していえば，それぞれの病院で 1 日ごとに生まれる赤ちゃんが標本であり，それぞれの病院で生まれるすべての赤ちゃんが母集団である。問題に，何年間にもわたって生まれる赤ちゃんの男の子の比率は約 50%であるという条件が含まれている。多くの対象者に，母集団における男の子の比率という性質は，標本における男の子の比率という性質と類似度が高いと判断したことにより，大数の法則の無視の傾向がみられるという説明である。

しかしながら，代表性ヒューリスティックのみで，統計判断における学習者の誤りを説明することは困難である。Kahneman *et al.*（1982）は，当初「大数の法則を理解していないために，標本の大きさが無視される。」という見解を示していた。しかし，その後の研究において，標本の大きさのみを変動させ，他の要因は統制された状況において，学習者は標本の大きさを考慮して判断することが示された。たとえば，400 人の標本調査と 1000 人の標本調査では，どちらの結果の方が信頼できるかという判断で，80%以上の対象者は標本の大きさが大きい方を選択したという研究結果がある（Bar-Hillel，1982）。Kahneman らも後になって，学習者は大数の法則を理解しているが，判断が求められる状況への適用に失敗しているという見解を示している（Kahneman and Tversky，1982）。そこで考えられるのが，どのような状況において，学習者は標本の大きさを考慮し，大数の法則を適用して適切判断をすることができるか，逆にどのような状況において，学習者は標本の大きさを無視し，代表性ヒューリスティックを発動させて誤判断するのかという問題である。

Evans & Dusoir（1977）は，コイン投げで試行回数という変数を 10 回，100 回，1000 回に変動させた条件のもと，様々な状況において対象者に判断を求めた。この問題では，2 つの実験結果に関して，表が出た回数と裏が出た回数の比が提示され，どちらの結果の方がコインに表が出やすい偏りがあることを示す強い証拠になるかを判断することが求められた。たとえば，「60:40＝7:3」が提示され，強い証拠になると考えた方に丸を付ける。もし，どちらも等しい程度の証拠になると考えたならば，等号に丸を付けるというものである。

（ⅰ）「比が一定・標本の大きさは変動」（たとえば「7:3＝700:300」）の状況

48 人中 22 人（45.8%）に等号に丸を付ける傾向がみられ，48 人中 15 人（31.3%）に標本の大きさが大きい方に丸を付ける傾向がみられた。

（ⅱ）「比が変動・標本の大きさは一定」（たとえば「60:40＝80:20」）の状況

48 人中 47 人（97.9%）に表が出た比率が高い方に丸を付ける傾向がみられた。

（ⅲ）「比が変動・標本の大きさも変動」（たとえば「700:300＝80:20」）の状況

48 人中 34 人（70.8%）に表が出た比率が高い方に丸を付ける傾向がみられた。また，（ⅰ）で標本の大きさが大きい方に丸を付ける傾向がみられた 15 人のうち 11 人（73.3%）に標本の大きさが大きい方に丸を付ける傾向がみられた。

「比が一定・標本の大きさは変動」の状況において，半数近い対象者（45.8%）に，標本の大きさを無視する傾向がみられた。この結果から，比が一定である情報が提示されると，多くの学習者は，標本の大きさを無視し，代表性ヒューリスティックを発動して，誤って判断すると考えられる。また，「比が変動・標本の大きさも変動」の状況において，7 割の対象者（70.8%）に，比の値の大小に基づいて，誤って判断する傾向がみられた。この結果から，比と標本の大きさに関して相反する情報が提示されると，多くの学習者は，標本の大きさよりも比の情報を優先して適用し，誤って判断すると考えられる。

以上から，今後の研究において，統計判断における誤りの要因を詳細に検討することとともに，教育的な視点から学習者の誤判断を修正する教授方略の構築が必要になるだろう。

標準偏差の大小判断に関する数学的検討

標準偏差の意味は，測定値と平均値との標準的な差を示す指標であるということである。測定値と平均値との標準的な差として，自然に考え出される指標は，偏差（Deviation）の平均値である。偏差の平均値とは，n 個の値からなるデータで，測定値と平均値との差を表す量の総和を n で割った値である。しかし，この値はどのようなデータにおいても必ず 0 になってしまうため，データの散らばりの指標としての機能を持たない。そこで，分散（Variance）という指標を用いるのである。分散とは，n 個の値からなるデータで，測定値と平均値との差を表す量の 2 乗の総和を n で割った値である。分散という指標も，測定値と平均値との差の 2 乗和の平均値であるため，設定する単位によっては大きな値になったり小さい値になったりしてしまい，直観的に測定値と平均値との差をとらえにくい。そこで，分散の平方根をとった値が標準偏差（Standard Deviation）である。標準偏差とは，n 個の値からなるデータで，測定値と平均値との差を表す量の 2 乗の総和を n で割った値の平方根である。標準偏差の大小を適切判断させるためには，標準偏差の求め方を指導するだけでは不十分であり，何が目的であり，どのような方法でつくられた指標なのかという標準偏差の意味を，学習者に理解させる必要があるだろう。

標準偏差 σ を計算で求めるときは，一般に次の公式を用いる。

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$$

標準偏差の大小判断に関する心理学的検討

標準偏差の大小判断問題に対して，多くの学習者に，ヒストグラムを誤読し，頻度の大小に基づいて判断してしまう傾向がみられた（Garfield and delMas，2005）。これは，ヒストグラムという視覚的な表現の特性が影響を与えていると考えられる。ヒストグラムは，一方の軸に変数の区間であ

る階級を割り当て，もう一方の軸に各階級の頻度をとるグラフである。それとともに，柱の面積が確率密度を表す面積グラフの 1 種である。近年の我が国の大学生が学習してきた学校教育において，ヒストグラムは数学の内容として扱われていなかった。そのため，標準偏差の大小判断における誤りは，ヒストグラムに関する知識の不十分さから生じるものなのか，あるいは正しい知識を持っていても生じさせる要因があるのかについて，調査を通して慎重に検討する必要があるだろう。

学習者がヒストグラムに関する知識を保持していることを確認した上で，分布の形状と標準偏差の意味を関連付けることは, 学習者の適切判断を促進させるために重要な教授内容であるといえる。

統計判断の誤りに関する調査

近年のわが国の大学生が保持している統計判断の誤りは，Kahneman らが指摘するものより多様に存在することが推測される。それは，次の理由による。近年のわが国の大学生が学習してきた学校数学において，統計の内容は希薄であった。中学校数学の内容から統計が削除され，高校数学では選択科目の一領域として扱われているに過ぎなかった（小口，2005)。そのため，近年のわが国の大学生にみられる統計判断の誤りは，アメリカ合衆国の大学生より多いと予想される。

本調査の目的は,「近年のわが国の大学生にみられる統計判断の誤りには，どのようなものがあるか。」という問いに答えることである。

調査の方法

１．調査の概要

学習セクションと事後調査からなる。学習セクションでは， 5 コマ（7.5 時間）の講義時間内に，高校数学 B「統計とコンピュータ」の内容を含んだ統計の基本的な概念を指導した。事後調査では，統計の基本的な概念に関する問題計 28 問を提示し，統計の概念と正答率の違いにより，問題を 3 段階に分類した。

２．対象者

私立大学文学部の「数学Ⅱ（統計学入門)」を受講した学生 68 名を調査対象にした。

３．手続き

(1)　学習セッション

次に挙げる統計の基本的な概念を指導した．高校数学 B「統計とコンピュータ」の内容と箱ひげ図，および相関と回帰の概念である。

①　分布の形状が対称か，歪みがあるか。

②　分布の中心傾向を測る代表値。

③　分布の散らばりを測る範囲と標準偏差。

④　箱ひげ図による表現と 5 数要約。

⑤　2 変量の相関と回帰直線。

(2) 事後調査

問 7 : 最近の研究で，調査対象者を，毎日摂取するビタミン E の量が違うレベルにしたグループに無作為に分けた。1 つのグループは，偽薬を受け取った。この研究は，ある期間内に特定のタイプのがんにどれくらいの人が発症するのかを調べるため，8 年間にわたって調査対象者を追跡した。次の記述の中で，この研究における無作為化の目的について，最もよく説明しているのはどの記述だろうか。

ア. 調査結果の精度を上げるため。

イ. すべての潜在的ながん患者が，この研究で等しく選ばれることを保証するため。

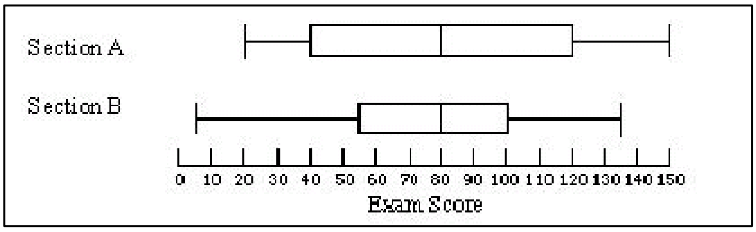
ウ. 標本抽出の偏りを減少させるため。

エ. 似た特性をもつ実験群をつくるため。

オ. 結果の歪みを防ぐため。

FIGURE 9 調査問題 7

次の 2 つの箱ひげ図は，同じ講義を受講している 2 つのクラスにおいて，期末試験の得点を表示している。



問 9 : どちらのデータの方が，30 点以下の学生の比率が高いですか。

ア. クラス A。

イ. クラス B。

ウ. 両方のクラスがほぼ等しい。

エ. どちらということはできない。

問 10 : どちらのクラスの方が，80 点以上の学生の比率が高いですか。

ア. クラス A。

イ. クラス B。

ウ. 両方のクラスがほぼ等しい。

エ. どちらということはできない。

FIGURE 10 調査問題 9・10

Garfield and delMas（2005）が提供している CAOS を利用し，学習者の誤判断を抽出していく。本調査では，CAOS の問題から，学習セクションで指導した内容で回答が可能な同一問題計 28 問を利用した。回答形式は選択式とし，標本抽出に関する問題 8 問，分布の形状に関する問題 8 問，分布の中心に関する問題 1 問，分布の散らばりに関する問題 4 問，箱ひげ図に関する問題 3 問，相関と回帰に関する問題 4 問をランダムに提示した。

調査の結果と考察

全 28 問の正答率の平均は 47.6%であり，範囲は 7.4%～92.6%であった。統計の概念と正答率により，問題を 3 段階に分類した結果を TABLE 1 に示す。

TABLE 1　正答率による問題の分類（数字は問題番号）

正答率 / 概念	**30%未満**	**30%以上 60%未満**	**60%以上**
標本抽出	7,16,26	13,24,27	23,25
分布の形状	6	2,22	1,3,4,5,11
分布の中心			12
分布の散らばり	17	14,15	18
箱ひげ図	9,10	8	
相関と回帰	28	20,21	19

（本稿は，下線の問題を分析対象とした。）

問 7（FIGURE 9）について，選択肢アは標本の大きさ，選択肢ウとオは標本の偏り，選択肢イとエは無作為化の目的に関する記述である。最も多かった誤反応は選択肢イで，その反応率は 32.4%であった。この結果から，無作為化の目的は，「母集団の傾向をとらえるために，母集団のすべての要素について，選ばれる確率が等しくなるように標本を抽出すること。」ととらえてはいるが，どの母集団の要素について，選ばれる確率が等しくなるようにするのかという無作為化の対象に関する誤判断があると考えられる。

問 9 と問 10（FIGURE 10）は，箱ひげ図の読み取りに関する問題であり，問 9 の正答は選択肢エである。問 9 では，多くの対象者に，箱ひげ図のひげの長さが長いグラス B（選択肢イ）の反応がみられた。問 10 の正答は選択肢ウである。問 10 では，多くの対象者に，箱ひげ図の箱の大きさ（箱

の右半分の大きさ）が大きく，データの上位 50%の範囲も広いクラス A（選択肢ア）の反応がみられた。これらの結果から，多くの対象者には，箱ひげ図の箱の面積やひげの長さは，データの頻度を表しているという誤読による誤判断があると考えられる。

また，問 16（キャンディー工場問題 : FIGURE 7）は，Kahneman らが指摘する「標本の大きさの無視」に対応した問題であり，正答は選択肢ウである。Karneman ら（1972）の産科病院問題に対する結果より，本調査のキャンディー問題に対する結果の方が正答率は低かった。

結論と今後の課題

本調査の目的は,「近年のわが国の大学生にみられる統計判断の誤りには，どのようなものがあるか。」という問いに答えることであった。その結果，次に挙げる誤判断が存在することがわかった。

分布の形状の誤判断

問 7（FIGURE 9）に対して，無作為化の対象として選ぶ母集団を間違えることによる誤判断があると考えられた。

頻度の誤判断

問 9 と問 10（FIGURE 10）に対して，箱ひげ図の箱の面積やひげの長さについて，面積の大きさやひげの長さがデータの頻度を表しているという誤読による誤判断があると考えられた。

標本の大きさの無視

問 16（FIGURE 7）に対して，Kahneman らが指摘した標本の大きさの無視というバイアスによる誤判断があると考えられた。

今後の課題として，相互に関連する 2 つの課題を考えている。

第 1 に,「代表性ヒューリスティック」を発動したためにみられる学習者の誤判断を修正するための教授方略を構築することである。「代表性ヒューリスティック」による誤判断の修正に関しては，先行研究において様々な教授方略が構築され，その効果を実証するための実験が行われてきた。しかしながら，事後調査において，産科病院問題に対する正答率は 50%前後の結果となった実験など，十分な効果が得られていない実態がある。それほど,「代表性ヒューリスティック」による誤判断を修正させることは困難であるといえる。今後，さらなる工夫を加えた教授方略を構築し，学習者の適切判断を促進させることが第 1 の課題である。

第 2 に,「代表性ヒューリスティック」以外の要因を調査し，あるとすればその要因を特定することである。先行研究で構築された教授方略は,「代表性ヒューリスティック」による誤判断の修正に十分な効果が得られていない実態がある。その理由として，修正が必要な要因が「代表性ヒューリスティック」ではない可能性も考えられる。今後，記述式問題などを利用して，誤判断の要因を詳細に検討することが第 2 の課題である。

謝辞

本研究は科研費【課題番号 22530990】の支援を受けている。

引用文献

Bar-Hillel, M. 1982. "Studies of representativeness", In: Kahneman, D., Slovic, P. & Tversky, A. (Eds.), *Judgment under uncertainty: Heuristics and biases*, 69-83, Cambridge University Press.

Evans, J. St. B. T. & Dusoir, A. E. 1977. "Proportionality and sample size as factors in intuitive statistical judgment", *Acta Psychologica*, **41**, 129-137.

Garfield, J. and delMas, B. 2005. *Comprehensive Assessment of Outcomes for a first course in Statistics.* University of Minnesota.

市川伸一. 1996. 『認知心理学４　思考』(東京大学出版会).

Kahneman, D., Slovic, P. & Tversky, A. 1982. *Judgments Under Uncertainty: Heuristics and Biases.* Cambridge University Press.

Kahneman, D. and Tversky, A. 1972. "Subjective probability: A judgment of representativeness", *Cognitive Psychology*, **3**, 430-454.

Kahneman, D. and Tversky, A. 1973. "On the psychology of prediction", *Psychological Review*, **80**, 237-251.

Kahneman, D. and Tversky, A. 1982. "On the study of statistical intuition", *Cognition*, **12**, 325-326.

麻柄啓一・工藤与志文・植松公威・進藤聡彦・立木徹.　2006.『学習者の誤った知識をどう修正するか』(東北大学出版会).

小口祐一. 2005.「高校統計教育カリキュラムへの要請とその対応に関する研究」『日本数学教育学会論文発表会論文集』 38, 403-408.

Tversky, A. and Kahneman, D. 1971. "Belief in the law of small numbers", *Psychological Bulletin*, **76**, 105-110.

Tversky, A. and Kahneman, D. 1974. "Judgments Under Uncertainty: Heuristics and Biases", *Science*, **185**, 1124-1131.