相澤裕介『統計処理に使うExcel活用法』カットシステム

# 理論的な要点（第2章）

文責：柳瀬陽介（2010/12/02）

# 第1章の復習

(a) なぜ平均値だけではデータのばらつき具合はわからないのだろう。

 (b) ヒストグラムとは何だろう。

(c) 分散とは何だろう。

(d) 分散の数式を説明してみよう。

(e) 標準偏差とは何だろう

 (f) z得点とは何だろう。数式でも説明してみよう。

 (g) 偏差値とは何だろう。数式でも説明してみよう。

# 第2章の要点

(a) 母集団 (population) と標本（sample）の違いは？なぜ全て母集団調査をしないのか？

・標本調査：教英学生30人の「標本調査」から「日本の英語学習者」についての結論を出すことは妥当か？ぎりぎり妥当と考えられる結論は何についてのものか（cf 有意抽出法）。

・無作為抽出法

・単純無作為抽出法（乱数表やコンピュータでの疑似乱数の使用）

・系統抽出法（番号のつけ方に周期性があれば偏りが生じる）

・多段抽出法（例、全国の市町村を無作為抽出、次にその市町村から個体を抽出など）

・層別抽出法（例、医者を予め診療科別などの等質のグループに分けておき、それぞれのグループから標本を抽出する）

参考：<http://aoki2.si.gunma-u.ac.jp/lecture/SampleSurvey/rem3.html>

(b) 標本調査での平均、分散、標準偏差は、母集団調査の時の出し方と一緒で、エクセル関数では、AVERAGE, VARP（もしくはVAR.P）, STDEVPを使う。これらは標本を一種の母集団とした上で、その平均、分散、標準偏差を出したものと考えることができる。

(c) しかし、多くの場合、調査が興味をもっているのは、ある特定の標本についての特徴ではなく、その標本が代表していると考えられる母集団の特徴である。したがって、標本調査のデータから母集団のデータ（母数 parameter）を推定する必要がある。

(d) 母平均は標本平均と同じとみなされる。

(e) 母分散は、分散の分母を「データの個数」ではなく「データの個数　–　1」にすることで推定される。このように推定された母分散を「不偏分散」と呼ぶ（pp. 54-57）。言い換えるならば、不偏分散とは、標本から得られたデータでもって、母集団の分散を推定して出した値である。

分散 = $\frac{Σ （値　–　平均値）^{2}　}{データの個数　}$

不偏分散 = $\frac{Σ （値　–　平均値）^{2}　}{データの個数 -1}$

不偏分散のエクセル関数はVARである。

注：Excel 2010から不偏分散の関数としてVAR.Sが追加された（SはSampleのSとして覚えよう）。ただし従来のVAR関数も使えるとのこと。

(f)　標準偏差とは、分散をルートで割ったものであった。標本から母集団の標準偏差を推定する場合は、不偏分散をルートで割ればよい。

エクセルでは、標本のデータから、母集団の標準偏差を推定する場合、STDEV関数を使う。

(g) 一般化可能性について

「研究の結果が母集団全体の傾向に沿った偏りのない一般的なものであると主張できる程度」を一般化可能性といいますが、この一般化可能性について判断する際に考慮すべき要因は、被験者の特性に関するものだけではありません。たとえば、児童・生徒の学業成績に対するある教授法の効果に関して検討することが目的の場合、通常、その効果が児童・生徒一般に対して有効であるか否かについてだけでなく、その教授法の中の種々のバリエーション、その教授法を適用する教科・単元、教授者の特性、教授者と児童・生徒の人間関係、学級集団の特性、従属変数の測定方法などの、さまざまな要因に関する多くの条件の差異を越えた（言い換えれば、これらの要因における限定条件なしの）一般的効果について論じようとしているはずです。しかし、一回の実験や調査においては、以上の諸要因に関して、ごく限られた条件下でしか検討を行うことはできません。たとえば一回一回の実験では、特定の教科の特定の単元のみについて、問題にしている教授法のその単元用のあるバージョンで、一人または少数の教師が授業を行い、ある限定された方法で従属変数である学業成績を測定して、検討を行うことになります。しかし、たとえば「新しく考え出されたある教授法は、従来の教授法に比べ、数学においては学業成績をより高めるが、社会においては差がみられない」などというように、以上の諸要因と教授法の間に学業成績に対して交互作用効果が存在する可能性があるため、ある限定された条件下で見いだされた効果が他の条件下でも同様に生じるという保証はありません。つまり、被験者に関して標本抽出の無作為性が現実には守られないことに目をつぶったとしても、一回限りの実験や調査で一般的な結論を下すことはとうてい無理なのです。言い換えれば、一回の実験や調査における仮説の支持は、あくまで仮説の例証（仮説の一支持例の提示）にしか過ぎず、仮説の完全な実証を意味するものではありません。結局、判断保留的態度のもとに研究を積み重ねることが重要なのです。また、分析結果について考察する際には、その結果を一般化する際の限定条件（自分の研究の被験者、独立変数の操作方法、従属変数の測定方法、剰余変数の統制方法、などの特徴）を、なるべく明確にしておくべきだと思います。なぜならば、このことは、過度の一般化を防ぐとともに、今後どのような新たな条件下で検討を重ねていく必要があるかについて考える際の指針になるからです。

統計的検定は、データを評価する（データが示している傾向が一般的にもあてはまるといえるか否かについて判断する）ための一つの基準でしかありません。そして、これまでに記してきたことから推察されるように、統計の力のみを借りて一般化可能性に関して完全な推論をすることは現実には不可能なのです。

世の中の法律や条令などの規則が非常識なことを自分勝手に行う不逞のやからを抑えるためにあるのと同様に、統計的検定は、本来、データの解釈の主観性を抑える（特に、各研究者が、偶然によっても十分に起こり得る程度の差でしかないものを過度に一般化して、軽率に各自の理論や仮説が正しいと主張してしまうのを防ぐ）ためにあるのだと思います。そして、統計的検定は、ある程度は実際にそのような有効な働きをしているといえるでしょう。しかし、その無批判な使用（ある一つの判断基準でしかないものへの過度の依存）が、かえって、統計的検定の結果が有意でありさえすれば意味のある現象についての一般化ができたと短絡的に結論づけてしまう、本来の意にそぐわない風潮をもたらしてしまっているようです。私たちは、いったん規則が作られると、どのような場合にもその規則に縛られた型にはまった見方をしてしまい、一つ一つの事柄に対して自分自身で柔軟に考えてみることをしなくなってしまいがちです。たとえば、規則に沿っていさえすれば常識的にみてどんなに不合理なことでも認めてしまったり、逆に規則に反していると本当に困っている人たちであるにもかかわらず手をさしのべなかったりすることが多々あります。そして、残念ながら、このような、いわゆる「お役所的な思考」が、自分自身の頭で考えることを本来もっとも要求されるはずの研究という場面においてもはびこってしまっているように思えてなりません」。

以上、吉田寿夫　(1998)『本当にわかりやすいすごく大切なことが書いてあるごく初歩の統計の本』北大路書房245-247ページより。

(h) 正規分布

これまで解説した統計手法は、母集団の分布が正規分布になる場合にのみ有効である。それでは「正規分布」(normal distribution)とは何だろう。

⇒教科書69ページを読もう。

偏差値とｚ得点は、データが正規分布していると仮定した上で、正規分布の特徴をデータが兼ね備えているはずだとして次のような推論をした。

偏差値40-60の間に全データの約68%が含まれる。

偏差値30-70の間に全データの約95%が含まれる。

偏差値20-80の間に全データの約99%が含まれる。

したがってあるデータが正規分布だと仮定されるなら、平均（母平均≒標本平均）と標準偏差（推定値）には次の関係があると推定される。

平均値±標準偏差の間に全データの約68%が含まれる。

平均値±標準偏差の2倍の間に全データの約95%が含まれる。

平均値±標準偏差の3倍の間に全データの約99%が含まれる。

しかし繰り返すが、これは母集団の分布が正規分布と仮定される場合にのみ有効な推論である。この仮定が無理のない過程であることを知るためにも、標本データをヒストグラム化して、直感的にも

(1) データが平均値を中心に左右対称に分布している

(2) 平均値の近辺はデータの頻度が高く、平均値から離れるほどデータの頻度が低くなっている

ことを確認しておくことは重要でしょう。

**※考えてみよう**:全国学力・学習状況調査(通称「全国学力テスト」)について

文部科学省では、全国的に子ども達の学力状況を把握する「全国学力・学習状況調査」を平成19年度から実施することとしました。
<http://www.mext.go.jp/a_menu/shotou/gakuryoku-chousa/index.htm>

その概要は次の通りです。

＜対象学年＞
○ 小学校第6学年、中学校第3学年の原則として全児童生徒を対象

＜実施教科＞
○ 教科に関する調査（国語、算数・数学）
・ 主として「知識」に関する調査
・ 主として「活用」に関する調査

○ 生活習慣や学習環境等に関する質問紙調査
・ 児童生徒に対する調査
・ 学校に対する調査

この調査の[意義](http://ja.wikipedia.org/wiki/%E5%85%A8%E5%9B%BD%E5%AD%A6%E5%8A%9B%E3%83%BB%E5%AD%A6%E7%BF%92%E7%8A%B6%E6%B3%81%E8%AA%BF%E6%9F%BB#.E6.84.8F.E7.BE.A9)や[問題点](http://ja.wikipedia.org/wiki/%E5%85%A8%E5%9B%BD%E5%AD%A6%E5%8A%9B%E3%83%BB%E5%AD%A6%E7%BF%92%E7%8A%B6%E6%B3%81%E8%AA%BF%E6%9F%BB#.E5.95.8F.E9.A1.8C.E7.82.B9)は様々に論じられています。

国際基督教大学(教育社会学)の藤田英典先生は、岩波書店の雑誌『[世界](http://www.iwanami.co.jp/sekai/)2009年1月号』の「有害無益な全国学力テスト--地域・学校の序列化と学力・学習の矮小化」という記事(232-240ページ)において、全国の学校・生徒を対象にした「悉皆」(しっかい)学力テストの実施により、大阪府・橋下知事、秋田県・寺田知事、鳥取県・平井知事などが市町村別あるいは学校別の成績開示を求め、その実施の有無により教育予算を差配する方針を打ち出したことを重く受けとめ、「教育予算を差配するといったことは、義務教育の機会平等を保障するという点でも、その質向上を公平に支援・促進するという点でも、あってはならないことである」(234ページ)と批判しています。(藤田先生はその他にもテスト学力という一元的な物差し (尺度) で子どもを評価する傾向が高まること、さらにそれにより学校・地域の序列化・格差化が進行することなどに対しても懸念を表明しています)。

しかし悉皆調査で事実上全ての学校の児童・生徒の成績情報が得られた以上、その情報を開示せよという要求はこれからも出てくることかと考えられます。

ここで悉皆テストという母集団調査(全体調査)は、適切な調査方法だったのかを、授業で学んだ**母集団調査と標本調査の違い**から検討したいと思います。

全国学力・学習状況調査の目的は次の通りです(注)。

＜調査の目的＞

○ 国が全国的な義務教育の機会均等とその水準の維持向上の観点から各地域における児童生徒の学力・学習状況をきめ細かく把握・分析することにより、教育及び教育施策の成果と課題を検証し、その改善を図る。

○ 各教育委員会、学校等が全国的な状況との関係において自らの教育及び教育施策の成果と課題を把握し、その改善を図るとともに、そのような取組を通じて、教育に関する継続的な検証改善サイクルを確立する。

○ 各学校が各児童生徒の学力や学習状況を把握し、児童生徒への教育指導や学習状況の改善等に役立てる。
<http://www.mext.go.jp/a_menu/shotou/gakuryoku-chousa/zenkoku/07032809.htm>

ここでは上記の表現を受け、同調査の目的を

(1) 各地域の学力・学習状況を把握する

(2) 各学校の学力・学習状況を把握する

とまとめます。

さて、ここで問題です。

悉皆調査という母集団調査・全数調査は、上記の目的(1)を達成するために必要なことですか? 目的(2)の達成のためにはどうですか? (時間があれば目的(1)と(2)それぞれの意味と波及効果についても考えてみましょう)。

ちなみに全国学力・学習状況調査の費用は、上記の藤田先生の記事によりますと、初年度は実施業者への委託費60億円以上に諸経費を加えて100億円を越えており、二年度も60億円以上かかっているそうです。また労力面でも大学入試センター試験並みの厳格な一斉実施方法のために莫大な労力がかかっているそうです。これは母集団調査として同調査を行なったための金額と労力だと考えられます。

(注) 全国学力・学習状況調査実施決定の際の文部科学大臣である中山成彬氏は記者会見で次のように述べています。

　全国的な学力調査につきましては、昨年11月に私が発表しました「甦れ、日本！」の中で、その実施について提案したところであります。本年6月21日に閣議決定されました「経済財政運営と構造改革に関する基本方針2005について」や、中央教育審議会義務教育特別部会の審議経過報告などにおいても、その実施の方向性が既に示されているところでございます。これらを受けまして、平成18年度概算要求に盛り込むべく準備を進めた結果、各学校段階の最終学年の小学校6年生と中学校3年生の国語、算数・数学について、全児童生徒が参加できる規模で平成19年度に調査を実施することとしたいと考えております。私といたしましては、全ての学校に対して、児童生徒の学習到達度・理解度を把握し検証する機会を提供することが重要であると考えておりまして、調査が円滑に実施されるように引き続き努力してまいりたいと考えております。
<http://www.mext.go.jp/b_menu/daijin/05090201.htm>

しかし一方で、中山氏は同調査を「日教組の強いところは学力が低いのではないかと思ったから」実施したと述べてもいます(朝日新聞2008年9月26日)。仮にこれが中山氏の真意だったとして、全国学力・学習状況調査は中山氏の目的を達成するために適切な方法だったといえるでしょうか?

**追記**

このテストは2009年での民主党による政権で見直しが行なわれ､悉皆調査ではなくなる予定です｡

**追追記 (2013/1/29)**

安倍政権は2013年1月29日に決定した新年度予算案で、この「全国学力テスト」を新年度以降、毎年すべての公立の小中学校で実施することにし、54億円を計上しました。（NHK報道）

\*\*\*\*\*\*\*\*\*\*

(i) 信頼区間

さきほど(d) で、「母平均は標本平均と同じとみなされる」と書きましたが、もちろんこれはそのように「みなされる」つまりは「推定することにする」ということです。「推定」なのですから、もちろん外れる場合があります。というより、厳密に「母平均 ＝ 標本平均」である確率は高くなく、たいていの場合、母平均は標本平均よりいくらかズレています。しかしそのズレ具合は、大きくズレる確率は小さくズレる確率よりも低いと考えられます。

このズレ具合を統計的に推定して、「まあズレたとしても、母平均はだいたいこの範囲にあるだろう」と算出した区間を「信頼区間」(Confidence Interval: CI)と呼びます。

「平均値の95%信頼区間が○○~△△である」というのは、○○~△△の範囲に母平均が95%で含まれるということです。

ということは、母平均が「○○以下または△△以上」である確率は5%ある（＝危険率）ということですが、この危険率を0%（つまり母平均の推定が絶対に正しい）となると、平均値の信頼区間は無限小~無限大となってしまいますから、危険率は通常5%にしておきます（高い精度の推定が必要な場合は、信頼区間を99%で設定し危険率を1%にする場合もあります）。

信頼区間の○○~△△は、標本平均からある値を引いたもの（○○）と、同じその値を足したもの（△△）で算出します。その値とは「t × 標準誤差」です。「t」は後で説明しますので、先に「標準誤差」(Standard Error: SE)を説明します。

「標準誤差」(Standard Error: SE)という用語は、前に習った標準偏差(Standard Deviation: SD)と似ていますね。実際、これら二つの用語は関連しています。母集団の散らばりが標準偏差でしたが、**「標準誤差」とは「標本統計量の標準偏差」です**。

標本統計量とは、標本から得られる母数に関する推定値です。(e) では、「母分散は、分散の分母を「データの個数」ではなく「データの個数　–　1」にすることで推定される」としましたが、もちろんこれも推定ですからその推定値である確率はある程度高いけど、それがドンピシャで当たっている（点推定）であるわけではなく、まあその推定値を中心としたある一定のばらつき具合の範囲の中に母分散はあるだろう、ということを意味しています。標本統計量におけるこのばらつき具合を「標準誤差」といいます。生のデータのばらつき具合でしたら標準偏差と言うのですが、標本統計量でのばらつき具合は「標準誤差」と言うのです。

さてこんどはtの値です。このtの値とは「t分布」の値です。これは次の三章でも学びますが、これは「データが正規分布をすると仮定されるが、正規分布のために必要な母数が得られない場合に、正規分布の代わりに推定する分布」のことです。非常に乱暴な言い方をしますと「正規分布の親戚」みたいなものです。

さて、信頼区間の○○~△△は、標本平均から「t × 標準誤差」を引いたもの（○○）と、「t ｘ 標準誤差」を足したもの（△△）で算出するものでした。

つまり

信頼区間＝標本平均 ± （t × 標準誤差）

です。

標準誤差は

$$t ×\sqrt{\frac{不偏分散}{サンプルデータの個数}}$$

ですから、

信頼区間 ＝ 標本平均 ± $\left(t ×\sqrt{\frac{不偏分散}{サンプルデータの個数}}\right)$

となります。

エクセルの関数では、

tの値は「=TINV(危険率,自由度)」（注）

ルートは「=SQRT(値)」

不偏分散は「VAR（セル範囲）」もしくは「VAR.S（セル範囲）」

ですので、これで信頼区間は算出できるはずです。（詳しくは教科書54-64ページをご覧ください）。

（注）Excel 2010からtの値を求める関数として

＝T.INV.2T（危険率、自由度）

が導入されましたが、これまでの「=TINV(危険率,自由度)」も使えます。なお危険率は5%なら「0.05」と入力してください。自由度はt分布の場合「データ数 – 1」です。