茨城キリスト教大学紀要第42号　287
自然科学　p.287～308

教育統計学の講義研究（１報）

土　井　　　努*

はじめに

本稿は筆者が1997年から受持っている教育統計学の授業についての概要をまとめたものであるが、文科系の学生が対象であるため、理論的な順序より理解性を優先して、統計学を差と相関の二概念からとらえた。その結果、講義においては、①二集団について特性の差が有意であるか否かを学ぶ事と、②二種類のデータに関する関連性を学ぶ事とに分けた。

①については、まず集団内のばらつきが問題となるので、これを実感するため、教室における的当て実験を行い、そして、ばらつきの中から集団の差を発見する問題へと展開している。②については、初歩的な統計学の教材としては、あまり例がないが、アンケート調査の分析を念頭に入れて、得点などの数値を用いた従来の相関に加えて、数値の代わり順位、属性などを用いた新しい相関の紹介を行っている。

統計学は一般に知られているように、教育の分野だけでなく、およそデータにばらつきの存在する分野を対象に、ばらつきの中から何かの傾向を読み取る目的で利用されている。しかし敬遠されがちな分野となる理由に、数式の取扱い以上に、現実の問題を統計の概念にモデル化する段階において困難がともなう事が考えられる。統計ソフトが普及した現在、数式の取扱いは主な問題とはならず、モデル化と結果の解釈とに問題点が集約されると思われる。

これらの点を踏まえて、授業では、まずデータを直感的に捉え、モデル化へのヒントとしている。また必要に応じてExcel上で筆者が作成したソフトを利用し、教育用の規模に合った感覚が得られるように試している。最後に各自の例題作成をテーマ毎のレポート課題としているが、この中から典型的な例、モデル化に注意を要する例などを紹介したい。

1．データのばらつきと集団の差

1.1　ダーツ実験から正規分布へ

1.1 a　目的

目標物に対して何かを当てようとするとき、一般にはどうしても誤差が生じる。このような誤差はごく普通に身の回りに存在するものであるが、このばらつき具合を実験的に知

*茨城キリスト教大学　文学部　非常勤講師

る事を最初の目的としている。実験は教室において各自が的当て（ダーツ）を行い、これを通じて誤差のばらつきは正規分布と呼ばれる数学モデルによって近似できる事、および、ばらつきの大小は標準偏差と呼ばれる指標で表される事を学ぶ。次の目的は正規分布の知識を基に、変数がある特定の値をとったときとその発生確率の関係を学ぶ事であるが、これは教育分野において得点と順位との関係に対応付けする事が出来、応用のための基礎概念となっている。

1.1ｂ　実験の手順

ダーツ実験は次の手順で行っている。

ａ．方眼用紙（５ミリ罫）の中央に、幅５ミリの十字を描き、これを狙うべき目標とする。

ｂ．床の上にこの用紙を置き、立った姿勢で鉛筆を持ち、目の高さから落とす。

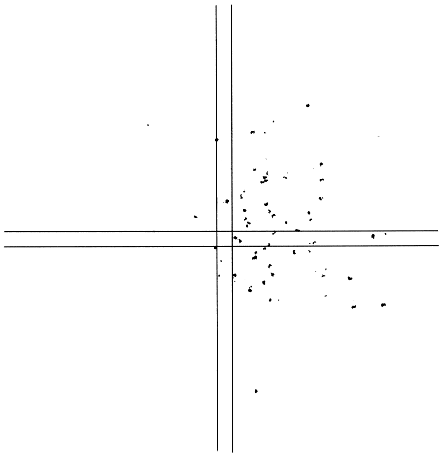
ｃ．この操作を約50回、精神統一して無心に行う。これは無作為性を保つ為である。

ｄ．集計は目標から５ミリのずれ量を１単位として、縦方向、横方向別々に行う。

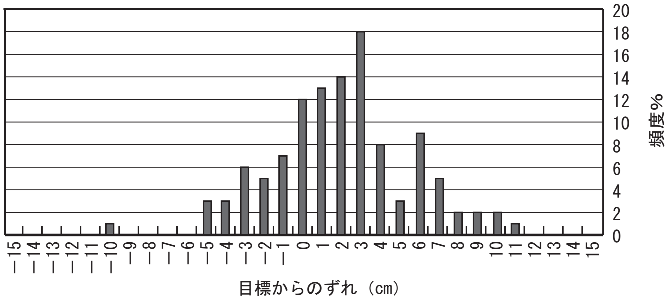
ｅ．最後に縦横合わせた集計を行い、頻度分布図を描く。

1.1ｃ　実験結果と正規近似（図表1.1　1.2　1.3）

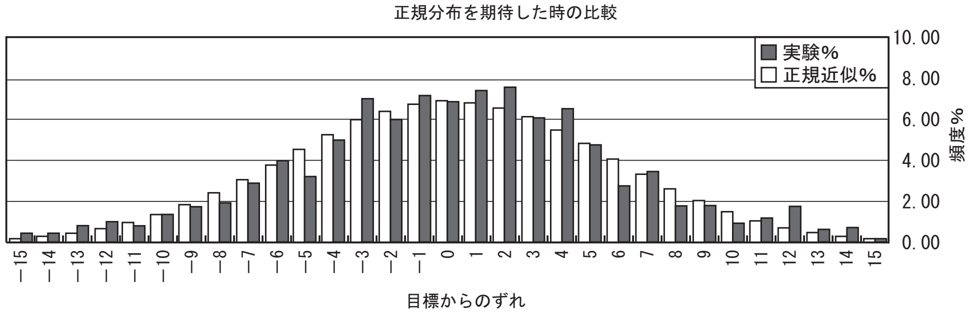
各自の実験結果の一例を図表1.1に、この頻度分布を図表1.2に示す。このように、各自の狙いの偏り具合、ばらつき具合は様々であることがうかがえる。一方、全員の実験結果を合わせた分布を図表1.3に示すが、このように誤差の分布は多数のデータを集めれば、正規分布によって近似できることがわかる。



図表1.1　ダーツ実験の１例



図表1.2　各自の実験結果１例



図表1.3　全員の実験結果

1.2　正規分布の利用

一般にある科目の得点と、順位との関係など、ばらつきのある多くの分析対象は正規分布で近似可能な事が知られている。そこで、例えば得点または順位の片方を知って、もう片方を推定する方法を学ぶ。つぎに各自が考えた応用例をレポートから紹介する。

1.2ａ　標準偏差に対する解釈

標準偏差は例えば１次関数のｙ切辺はこの場所と言うようには、直接示す事ができないので、理解を難しくしている。しかし多くの分析対象は正規近似可能であるので、この場合は次のように言う事ができる。

ａ．平均値±標準偏差　の範囲に含まれるサンプルの割合は、全体の68%すなわち約2/3である。これを人々の意見に対応させて考えれば、重要な提案を成立させる為の大方の意見ということになり、社会学的な意味付けを標準偏差に持たせる事ができる。

ｂ．実用的な分析の範囲は、平均値±２×標準偏差　以内の数値が対象となるので、これより多少余裕を持った範囲を想定する場合が多い。

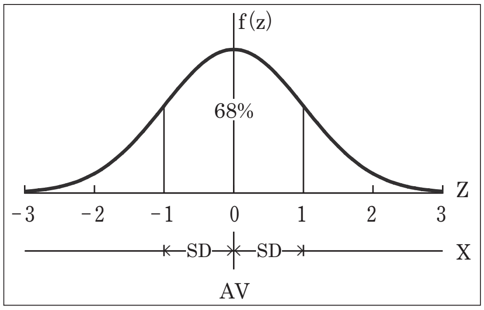
1.2ｂ　正規分布の標準化

各種の実用問題に対して、正規分布を適用するため、変数の標準化を考える。分析対象と考える集合全体を母集団と呼び、これは十分な個数を持ったデータの集まりと想定する。

これに対して、ある具体的な取得データをサンプルと呼び、分析の目的は想定した母集団中において、このサンプルがどのような位置に存在するか、または、サンプルの発生確率はどの程度かなどを推定する事にある。応用のため母集団についてつぎの仮定が必要である。

a．分析対象データは、正規分布で近似可能とする。
b．分析対象すなわち母集団の平均値と標準偏差は既知とし、他の情報は無くても良い。

このとき次の変換式で表されるZは、平均値０、標準偏差１の正規分布に従う。Zに関する解釈は、この式によって現実のデータXを、正規変数Zの世界に、変換したものと見ることができる。なを、正規分布の形を知るために、密度関数f(z)を図表1.4に示す。



図表1.4　正規分布

変換式：$Z=(X-AV)/SD$

$f(z)=(\sqrt{(2\pi)})^{-1}\exp(-Z^2/2)$

ここで、

Z　　：標準化された正規変数
X　　：あるサンプルデータの値
AV　：母集団の平均値
SD　：母集団の標準偏差
f(z)　：正規密度関数

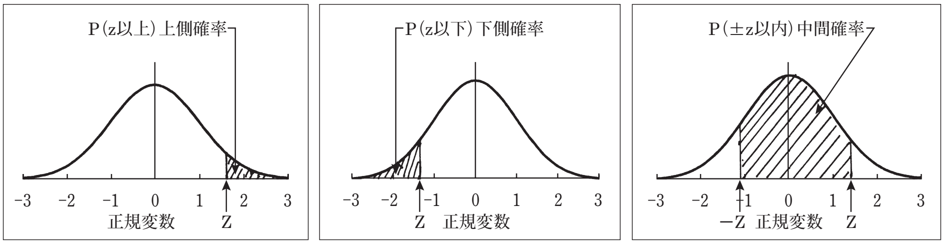
1.2c　正規分布における発生確率（図表1.5）

ここでは正規変数Zがある値をとったとき、これに対応した発生確率P(z)について考える。理解の為、これを次の３種類に分けて考える。

P(z以上)：上側確率、すなわち正規変数が、Z以上となる確率で、上からの順位を表す
P(z以下)：下側確率、すなわち正規変数が、Z以下となる確率で、下からの順位を表す
P(±z以内)：中間確率、すなわち正規変数が−Z以上＋Z以下となる確率

ある正規変数値Zから発生確率P(z)を求める、あるいはその逆を行う為には、次のいずれかの方法によるのが一般的である。

a．数表によって、ZとP(z)との対応を求める。
b．Excel関数を用いる。P(z以下)＝normsdist(z)、Z＝normsinv(P(z以下))
　なを、発生確率P(z)は正規関数f(z)を該当するZの範囲で積分したものと決める。



図表1.5　正規分布における発生確率

1.2ｄ　正規分布の応用例

正規分布の応用では次のＡ、Ｂタイプ両方について上側、下側、中間確立各々の例を用いる。

Ａタイプ：発生確率をある意図から指定し、その様に成るデータ値Xを求める問題。
すなわち、P(z以上)など→Z→X なる手順。

Ｂタイプ：あるサンプルデータXを得たとき、そのような発生確率を求める問題。
すなわち、X→Z→P(z以上)などの手順。

図表1.5a　正規分布の応用例一覧

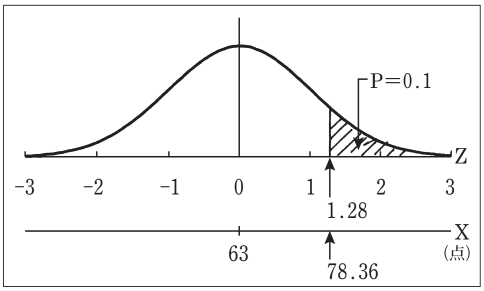
	Ａタイプ　　Ｂタイプ
上側確率	←例１　県統一テスト　→
下側確率	←例２　許容待ち時間　→
中間確率	←例３　携帯電話の料金　→

例１：上側確率

県統一テストにおける英語の得点は、平均AV＝63点、標準偏差SD＝12点と発表されたと想定し、他の情報は不明であると仮定する。ただし得点は正規分布に従う事がわかっているとする。(付録１)

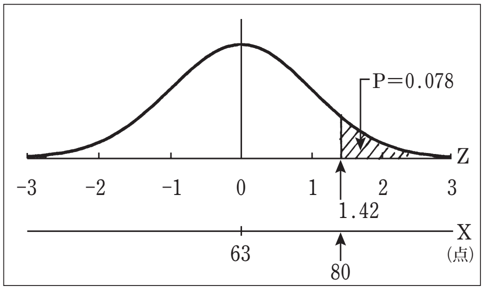
Ａタイプ：ある学校では100人中、10位までの生徒を評価Ａにすると決めるとき、そのような生徒の得点は何点以上であるか？(図表1.6)

解：100人中10位までとは、上側10％に対応するので、
正規分布表より、P(Z以上)＝0.1
となるようなZは、Z＝1.28
変換式より、1.28＝(X−63)／12
X＝78.36
よって約78点以上の生徒が上位１割に該当と推定。



図表1.6　Ａタイプの解

Ｂタイプ：この学校では別案として、得点が80点以上の生徒を評価Aにしようと思う。このような生徒は100人中、何位以上と推定されるか？（図表1.7）



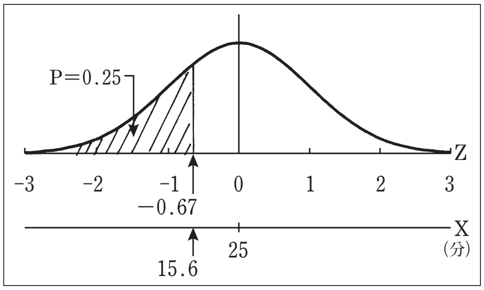
図表1.7　Ｂタイプの解

解：100人中の順位を発生確率で表せば
P(z以上)となるので、変換式より、
Z＝(80−63)／12＝1.42
正規分布表より、
P(1.42以上)＝0.078
これを順位で表現すれば、100人中７、８位より上と推定される。

例２：下側確率

待ち合わせにおいて、20代の人が許容できる待時間は、平均AV＝25分、標準偏差SD＝14分との調査結果があったと仮定し、他の情報は発表されていないとするが、ここでは、待ち時間は正規近似できるとする。(付録２)

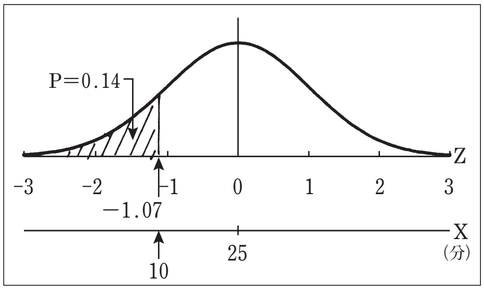
Ａタイプ：許容待ち時間が短い方の人を、短気とし、全体の25％を占める人々であると決めたとき、このような許容待ち時間は何分以下となるか？（図表1.8）



図表1.8　Ａタイプの解

解：正規変数Zは小さい方の値となる。
すなわち、P(z以下)＝0.25
となるようなZは、Z＝−0.67
変換式より−0.67＝(X−25)／14　X＝15.6
許容待ち時間が約16分以下の人は短気と推定。

Ｂタイプ：別の考えより、許容待ち時間10分以下の人を、短気と決めるとき、この調査では何％の人が該当するか？（図表1.9）



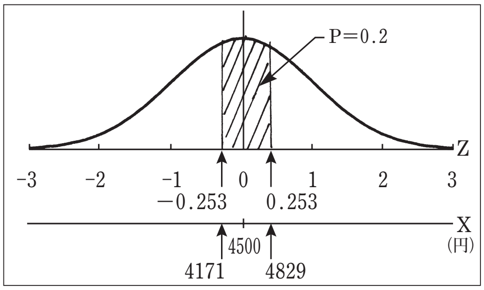
図表1.9　Ｂタイプの解

解：10分に相当する正規変数は、
変換式より、
Z＝(10−25)／14＝−1.07
このZ以下の発生確率は正規分布表より、P(−1.07以下)＝0.142
よって、短気な人は約14％いると推定できる。

例３：中間確率

ある携帯電話会社が、中学生の通話料金について調べたところ、一ヶ月の平均AV＝4500円、標準偏差SD＝1300円であったと仮定する。（付録３）

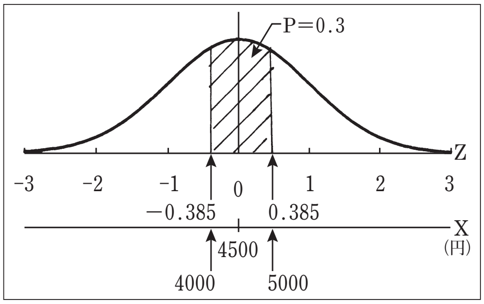
Ａタイプ：ある家庭では長く話したい子供側と、料金を節約したい親側の妥協案として平均を挟んだ通話料金の範囲で、全体の20％の中学生が該当するような金額を認める事にした。このような通話料金の範囲はいくらか？（図表1.10）



図表1.0　Ａタイプの解

解：対応する正規変数を±Ｚとすると、
P(±z以内)＝0.2となるようなZは、
正規分布表より２×(0.5－P(Z以上))＝0.2を解いて、Z＝0.253
変換式より0.253＝(X－4500)／1300　X＝4829
また、－0.253＝(X－4500)／1300　X＝4171
これより料金範囲は、4171円～4829円と推定。

Ｂタイプ：この家庭では、前述の料金範囲を、4000円～5000円と切りの良い数字に決定した。このような料金範囲に該当する中学生は全体の何％いると推定できるか？（図表1.11）



図表1.11　Ｂタイプの解

解：通話料金に対応する正規変数は変換式より
Z＝(4000－4500)／1300＝－0.385
またZ＝(5000－4500)／1300＝0.385
正規分布表より、２×(0.5－P(0.385以上))を解いて、P(±0.385以内)≒0.30
よって中学生の30％がこの範囲の通話料金に該当すると推定される。

1.2ｅ　レポートから

正規分布の理解を確実にするため、前期末にレポートとして各自の例題作りを課題にしているが、以下ではこれより選んだ、典型的な、あるいは問題を含んだ応用例を紹介したい。

母集団全体を分類する型

・コンタクト使用者が１日にさす目薬の回数は、多数の人を調査すれば、正規分布となるだろう。回数が多い人から順に乾燥目、健康目、涙目などの分類が考えられる。

・同様な例として、20代の人が１日に鏡を見る回数を調査したとする。この回数を変数とした分布が考えられ、この変数は、おしゃれに関する関心度の指標となるかもしれない。
・大学のある講義について、年間出席回数あるいは居眠りした授業回数を変数としてみる。これらの回数に該当する学生の分布を考え、一見まじめタイプ、普通タイプ、一発勝負タイプなどと分類する事が考えられる。さらに曜日、時限などによる分布の違いも面白い。
・外食するときの食事代を月給に対する割合で考え、これに関するサラリーマン人数の分布を考えると、食事代の多い方から、消費派、普通派、堅実派などと分類できるだろう。

上の例に表れた回数は、あまり大きくない整数と考えられる。正規近似は元来、連続変数を対象としているので、分析すべき整数は連続的でないことに注意を要するだろう。

特異データを決める型

次の例のように、ある１つのデータが、特異かを問う問題は、「特異」が意味するところを慣例に従い、全体の５％しか占めないような確率の少ない場合と、解釈する事にする。

・臨床検査一般について、検査値は、多数の人々を対象とすることにより、正規変数と考えられているので、不健康群に区分された人の検査値を、例え症状が出ていなくても、発病の予測などに使う事が出来る。
・散歩の好きな小太郎おじさん。普段は平均50分、標準偏差７分で散歩から帰ってきますが、ある雨の日、いつもと同じつもりが42分でした。この日は特別に短い方といえますか。
・化学の実験で、塩酸と水酸化ナトリウムの中和を考える。薬品を計るときの誤差によって、中和のPH値が変化する。このPHを変数としたそのような生徒の割合の分布は、例年のデータを蓄積すれば、正規近似できるだろう。このとき、実験ミスとは、分布の両裾、すなわち酸、アルカリ度が一定以上となると考えられる。そのような生徒の割合をあらかじめ知っておけば、授業時間などの予測に役立つかも。

一般データを決める型

・結婚年齢を変数とするデータが正規分布したとする。適齢期なる言葉の定義を決める為、平均を挟んで全体の2/3の人が結婚する年齢の範囲を適齢期とすることを考えた。なを、全体の2/3とは変数の範囲が平均±標準偏差（歳）に相当する。
・スーパーで主婦が買い物に要する時間を変数としたときの、主婦人数の分布を考える。普通の主婦（平均を挟んで全体の2/3を占める人とする）の買い物に要する時間の範囲は？
・金銭が絡む話題について、立場の違いで〜円以上、〜円以下となるような分析例：下宿生活学生の毎月の食費データがあり、正規近似できたとする。親が退職した家庭について「学生本人は死にそうなので〜円以上は仕送って欲しい、一方、年金生活の親にとって仕送りは〜円以下にして欲しい」と主張したとする。このような場合、両者の主張を

取り入れた食費は、何％の人をカバーできるかを調べることができ、これが一般性を持てば説得力が加わるだろう。

1.3　集団の差と有意性検定

これまでは１つの母集団が正規分布に従うときの問題を扱ってきたが、ここでは２つの母集団、例えば、初心者対ベテランドライバーなどを考え、集団に特性の差があるか否かを調べる。これは対象データにばらつきがあって、サンプル平均の差に意味があるか否かが即断できないような場合に適用される。サンプル数は多いほど望ましいが、費用、時間などの実用的な見地から、約10前後を想定する場合が多い。

1.3ａ　検定の考え方

検定の基礎概念を次に示したが、これは統計というより、通常の論理的な展開に基づいているとみなせるだろう。

①仮説「集団の平均に有意差は無い」を設ける。
　帰無仮説とも呼ばれ、背理法的な意味において、否定される事に存在意義がある。
②対立仮説「集団の平均に有意差があり、片方が大きい」を設ける。
　多くの場合、対立仮説が分析者の意図となる。すなわち対象を決定するとき、何らかの一般通念または事前知識があり、これを統計分析するのが意図となろう。例えば、ベテランドライバーの方が、初心者より走行速度が高いだろう、などである。なを、「集団に差はあるが、どちらが大きいか不明」の場合は③における確率を１/２にする。
③サンプルによる現実のデータは、①の仮説の下では、低い確率（通常５％）でしか起こらない事が示されたとき、
④ごく普通のサンプルにおいて、低い確率の現象が起こった理由は、①における仮説が間違っていたと判断して、これを捨てる。さらに②の対立仮説を採用する事、すなわち事前知識をこのデータから有意であると認める。
⑤一方、③の代わり、実データの起こる確率が、それほど低くない事（通常５％より大）が示されたとき、
⑥結論の表現は「２集団の平均値に有意差があるとは言えないと、このデータからは判断される」となる。しかし、目的は対立仮説が成立して初めて、分析の意味がある場合が多いので、サンプル数を増加させるなどの方法を考える。
⑦実用上、集団の平均に差が無いという事が目的の場合は、５％より十分大きい発生確率をサンプルデータから導く必要があるが、厳密には背理法的な論理に従っていないので、注意を要する。

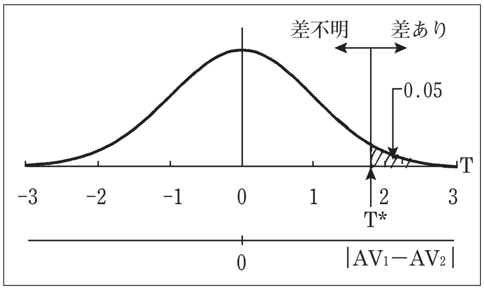
1.3ｂ　基礎理論（図表1.12）

２つの異なる母集団から複数のサンプルを得たとする。このとき平均値および標準偏差もサンプルにより変化する確率変数となるが、これを次表のように表す。また、母集団の平均、標準偏差などは未知とするが、ばらつきは２集団間で等しいとする。なを、以下で

は両集団からのサンプル個数は、簡単のため等しくしているが、そうでない時にも、拡張可能である。

	母集団平均	サンプル平均	サンプル標準偏差	サンプル個数
母集団 1	M1	AV1	SD1	N
母集団 2	M2	AV2	SD2	N

・まず、仮説と、対立仮説を1.3 a に従って立てる
　仮説：2つの母集団の平均に差が無い。
　対立仮説：片方の母集団平均の方が大きいとする。（以下の分析では絶対値をとるので、どちらの母集団でも同じ。）



図表1.12　有意差ありとする危険率

・次にサンプル平均のばらつきを推定する必要があるが、これは次式の合成標準偏差SDによる。（付録 4 ）
　$SD^2=((SD1)^2+(SD2)^2)/N$

・一方、差の統計量Tは、両母集団の平均に差が無いとの仮説を利用して、次のように計算される。ここで仮説が役立っている事は重要である。
　差の統計量T　：$T=|AV1-AV2|/SD$

・このTは自由度2N－2のt分布に従う事がわかっているので、サンプルデータの発生確率P(T以上)を求める。方法は、excelの統計関数tdistによるのが便利である。
　P(T以上)≦0.05のとき　―――平均の差は有意。
　P(T以上)＞0.05のとき　―――平均の差は有意とは言えない。

・数表によるときは、t分布において上側 5 %となるときの統計量T*を求め、差の統計量を変形した次式に代入する
　サンプル平均の差：$|AV1-AV2|=T\times SD$

・T*を代入した上式の右辺T*×SDは実単位上の限界値なので、次の結論となる。
　$|AV1-AV2|\geqq T^*\times SD$　―――平均値の差は有意。
　$|AV1-AV2|<T^*\times SD$　―――平均値の差は有意とは言えない。

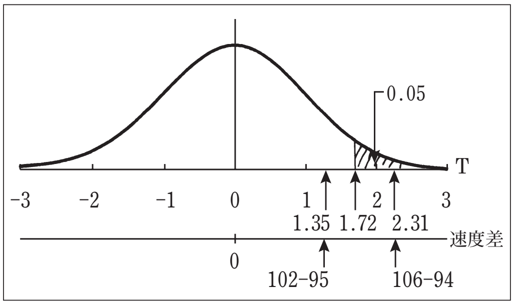
1.3 c　例題による有意差の検定（図表1.13）

ここでは 2 つの母集団として、自動車の運転についての初心者集団とベテラン集団とを考える。高速道路を通常の交通量のとき適当な時間、運転してもらい、その間の平均速度

を特性値として記録した結果が、次の表であったとする。

	母集団平均	ケース１の平均	ケース２の平均	標準偏差	サンプル数
初心者	M1	94km/h	95km/h	10km/h	12
ベテラン	M2	106km/h	102km/h	15km/h	12

・仮説､対立仮設を次のように考える。
仮説：ベテランと初心者間に速度差は無い：M2－M1＝0
対立仮説：ベテランの方が平均速度が高い：M2－M1＞0



図表1.13　走行速度の差

・合成標準偏差SDを計算する。
$SD^2=(10^2+15^2)/12$よりSD＝5.20

・サンプルによる差の統計量Tをまず求める。次にTは自由度12＋12－２のt分布に従うので、この発生確率P(T以上)を計算する。
ケース１：T＝(106－94)／5.20＝2.31　P(T以上)＝0.015
ケース２：T＝(102－95)／5.20＝1.35　P(T以上)＝0.095

・結論は、発生確率が0.05より小さい場合は、集団平均の差が有意となるので、
ケース１：平均の差は有意である。すなわちこのデータからは、高速道路における走行速度について、ベテラン運転者の方が、初心者よりも平均的に高いといえる。
ケース２：平均の差は有意とならない。すなわちこのデータからはベテラン運転者の方が、初心者よりも走行速度が平均的に高いとは言えないので、さらにデータ数を増して検討する必要がある。

・数表による場合は、t分布表において、自由度12＋12－２、上側確率５％に対応するt統計量値は、T*＝1.72
実単位上の限界値は、T*×SD＝1.72×5.20＝8.94km/h

・サンプルによる平均の差と、実単位上の限界値とを比較して、④と同様の結論を得る。
ケース１：AV2－AV1＝106－94＝12＞8.94　－－－平均の差は有意。
ケース２：AV2－AV1＝102－95＝7 ＜8.94　－－－平均の差は有意とは言えない。

1.3d　レポートより

以下に、レポートの中から、集団平均の有意差に関するいくつかの例を掲げる。
・掛け算九九を１分間に出来る回数について、大学生対小学６年生で有意な差が出るか？
計算速度は個人差が大きいので、多めのサンプルが必要となるだろう。

・お餅いっぱいの正月休みがきた。休みを挟んだ２週間の体重変化は、増加するかを考える。これを分析するには個人差を打ち消す為、同一の人々を対象として２週間の体重変化を調べる事が望ましいが、このときは多少異なるモデル化が必要となる。
・資源の有効利用のため、自動販売機による紙コップ回収と返金制度を考える。返金が５円のときと10円のときとの回収率について有意差を論じる。業者にとっては回収率が上がった分、支出が増えるので、このバランス点が問題となるが、統計上の有意性が金銭的に意味あるかを検討するのは大きな課題であろう。また、平均値の差に代わって率の差となっているので、統計モデルが変化する。

以下は、対策や処理などの効果、効用を判断する一般的な例である。
・ある塾の生徒でセンタ試験の前日、遊んだ群と、勉強した群とで、合格率に有意な差が出るか、知りたいとする。ただし選んだ生徒達は似通った成績とする。この例も率の差を論じるものであるが、率を計算する為、サンプル数は多く必要となろう。

２．クロス表における変数間の関連性

クロス集計表とは、分割表とも呼ばれ、２種類の変量（例えば英語の成績Ａ、Ｂ、Ｃと、数学の成績Ａ、Ｂ、Ｃなど）があるとき、変量を組合せたクロス項目に該当する人数、件数などを集計したものであり、種々あるデータベースの中でも、最も直感的に理解しやすい集計方法といえる。特に単なる関連性でなく、因果関係を調べたいとき利用価値が高い。このような変量について関連性の有無と、その程度を学ぶ事は、アンケート調査の分析などに威力を発揮するものと思われ、ここ数年、講義において従来の統計に加えている。

クロス表における関連性について、従来の統計は、例えばテストにおける２科目間の得点などを対象とするものであるが、関連性を調べたい変量は、得点などの数値以外の場合として、国名など物事の名前や、成績A、B、Cなどの順序の場合もある。一般に変量は３種類に分けられ、物事の名前のとき名義尺度、順序を表すとき順序尺度、数値のとき間隔尺度と呼ばれているが、クロス表の変量としては、これら３種類を対象としている。

クロス表の分析においては、まず表の縦横となる変量間に何らかの関連性が見られるか否かの有意性検定を行うことの重要性を強調している。これは、関連性の程度を知るために、今日では、統計ソフトを用いて、容易に相関係数などの結果を得易くなっているが、この場合は、関連性が有意となって初めて、相関係数の意味が生じるなど、計算に欠かせない前提条件が重要だからである。

2.1　名義尺度における相関と関連性

クロス表の縦横変量が、名義尺度である場合を対象として、変量間の関連性検定、より狭義には因果関係が考えられる場合は有意性の検定を行う。これは分析の第一歩として重要であるが、抽象的で理解しにくいので、具体例によって手順を示し、またその中で、相

関係数についても同時に扱うことにしたい。

名義尺度における例題（主食とデザート）（図表2.1）

次の表は100人の主婦を対象として、３種類の異なる主食を摂った後、デザートの好みについてアンケート調査をしたという、想定例であり、表中の数値は、そのような主食とデザートの組み合わせが好ましいと答えた人数である。

A表：データ

	シャーベト	和菓子	ケーキ	
和食	10	8	15	33
エスニック	13	12	8	33
洋食	22	9	3	34
	45	29	26	100

B表：無相関

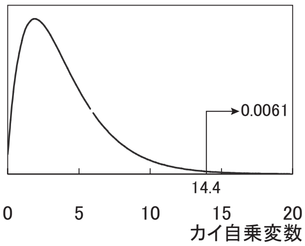
14.85	9.57	8.58
14.85	9.57	8.58
15.30	9.86	8.84

C表：カイ自乗

1.58	0.26	4.80
0.23	0.62	0.04
2.93	0.08	3.86

カイ自乗＝	14.4
Cramerの係数 r＝	0.268
関連性有意の危険率 P＝	0.0061

（カイ自乗検定値）



図表2.1　名義尺度の例：主食とデザートの好み

Ａ表：データ

クロス表の縦項目は、主食の種類（名義尺度）であり、この例では原因側と見る事ができ、横項目はデザートの種類（やはり名義尺度）であり、結果と見る事ができよう。また、主食の種類が原因側なので釣合いをとって調査人数を等くした。

2.1ａ　関連性の有意検定

①仮説として、「名義尺度から成る２変量は、独立である」すなわち関連性が無いとする。対立仮説として「２変量には関連性がある」とする。

例におけるＢ表では、

仮説を採用したとき、縦横の項目は独立となり、この例では、主食とデザートとに因果関係が無いことになる。表中の数値はこの仮定の下における、各セルの期待値を表している。

期待値＝(該当行の周辺和×該当列の周辺和）／総数

計算例：(上から２行目、左から３列目のセル要素）　8.58＝33×26／100

②クロス表のカイ自乗値を計算する。Ｃ表における各セルのカイ自乗値は、Ａ表のデータ値が、Ｂ表の期待値からどの程度隔たっているかを表したものである。

カイ自乗値＝(実データー期待値)^2／期待値

例におけるＣ表では、

計算例：(上から３行目、左から１列目のセル要素）　2.93＝（22−15.3)^2／15.3

表全体のカイ自乗値＝全セルのカイ自乗値の和＝14.4

③データから計算した②のカイ自乗値を用い、自由度＝（行数－１）×（列数－１）におけるカイ自乗分布の上側確率を、数表から求める。この上側確率は仮説を棄却し、縦横変量に関連性ありと結論するときの危険確率（P値）を表している。

危険確率(P値)≦５％(許容される危険確率)　ならば、危険確率が十分小さいと見做して仮説を棄却し、対立仮説を有意とする。すなわち縦横変量に関連性または因果関係ありと結論する。なを、この危険確率を、関連性に関する有意確率とも呼ぶが、この呼び方は、確率が大きい程、好しいかのような誤解を招きやすい。

例における結論：
カイ自乗分布における自由度＝（３－１）×（３－１）＝４
自由度４における②のカイ自乗値14．4に対する上側確率＝0.61％
すなわち、関連性有りとするときの危険確率＝0.61％＜許容される危険確率５％
これより主食の種類とデザートとは関連が無いという、仮説が棄却され、主食の種類によってデザートが左右されるという、対立仮説を有意とする。

2.1ｂ　特定セルと関連性

ｃ表において、カイ自乗値の大きいセルを調べる。

例）上から１行目、左から３列目のセル4.80：和食＆ケーキの組合わせ
実データ15人に対して、期待人数8.58人と、実データ人数の方が多いこと、および、カイ自乗値は4.80とセルの中では大きいことより、和食＆ケーキの組合せは非常に好まれる点で、関連性が強いことを示している。

例）上から３行目、左から３行目のセル3.86：洋食＆ケーキとの組合せ
実データ３人に対して、期待人数8.84人と、実データの方が少ないこと、およびカイ自乗値性が3.86とセルの中では大きいことより、洋食＆ケーキの組合せはあまり好まれない点で、関連が強いことを示している。

2.1ｃ　Cramerの属性相関係数

クロス表を構成する２変量が名義尺度であるとき、２変量の関連性は、属性相関と呼ばれており、数種類以上の相関係数が提案されているが、確定したものは無いのが現状である。講義では最も単純明快と思われるCramerの相関係数を取上げている。

Cramerの相関係数＝ √{クロス表のカイ自乗値／Ｎ／（ｔ－１）}
Ｎ：クロス表の全データ合計
ｔ：　行数、列数の小さい方

例）主食とデザートの例におけるCramerの相関係数は次のように計算される。
√{14.4／100／(３－１)}＝0.268

クロス表における縦横項目について、関連性の有意検定の結果、主食とデザートとに好みの因果関係が強いことがわかったが、Cramerなどの属性相関ははこの前提の下に計算された表全体としての関連指数と見る事ができる。つまり、関連性有りの検定を通過したので、属性相関の計算に意味が出たのである。逆に、関連性が無くとも、属性相関は計算可能であるが、その値の意味は根拠を失う。

Cramerの相関係数は、順序尺度、間隔尺度における他の相関係数と比較可能なように次の性質を持っている。

・相関係数は０から１の値をとる。
・相関係数が０のとき、縦横の名義尺度は独立であり、因果関係などの関連性は無い。
・相関係数＝１はクロス表の縦横項目が完全な関連状態にあることを表す。（付録５）
・名義尺度には順序が無いから、縦または横項目の並び替え順は自由である。
・上述の事から、関連性には比例、逆比例の概念は無くなり、単に結びつきの強さだけを表すため、相関係数は正の値のみとなる。

2.2　順序尺度と関連性

一般に変量が順序尺度の場合、データは数値ではないが、順に並べることは可能である。このとき、データ値に代わって、並べた順位の値を用いれば、一般の相関係数と全く同様の式によって、順位相関係数を定義する事が出来る。Spearmanの順位相関は、このようにして、２つの順序尺度間に定義された相関係数であり、比較的理解しやすいという点から、講義では順序尺度の関連性指標として取上げている。

2.2 a　Spearmanの順位相関の計算

クロス表においてSpearmanの順位相関係数を考えるときは、多少の注意が必要である。例えば縦変量A、B、Cなどの順位は１、２、３となりそうであるが、仮にAが１位から11位を占めた場合、平均は６位といったように、A、B、Cに平均順位を与えなければならない事がわかる。横方向についても同様である。クロス表は、セルの全個数が多くないので、逆に、同順位のデータ個数が、極めて多いのが特徴となる。

Spearmanの順位相関係数Rs

Rs＝縦横の共分散／（縦の標準偏差×横の標準偏差）

縦横の共分散＝$\Sigma\{N_{xy}\times(X-AV)\times(Y-AV)\}/(N-1)$

縦の標準偏差^2＝$\Sigma\{N_x\times(X-AV)^2\}/(N-1)$

横の標準偏差^2＝$\Sigma\{N_y\times(Y-AV)^2\}/(N-1)$

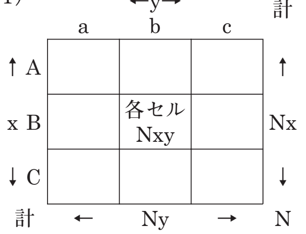
N_{xy}：各セル内のデータ個数

N_x：各セル内のデータ個数を横方向に合計

N_y：各セル内のデータ個数を縦方向に合計

X：　各縦項目A、B、Cなどの平均順位

Y：　各横項目a、b、cなどの平均順位



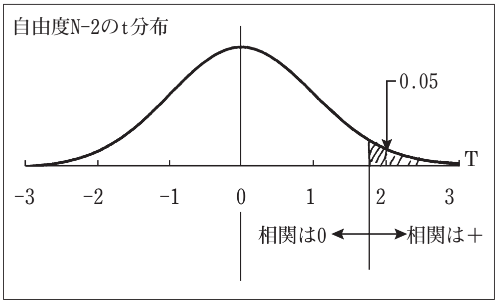
縦横項目が順序の３×３クロス表

AV: 全項目(全縦項目または全横項目)の平均順位≒N/2
N:　セル内の全個数

Spearmanの順位相関は、次の特徴を持っている。
・一般の相関係数（ピアソンの積率相関係数）において、変数値の代わり、変数の順位を用いたものなので、同一の計算式が適用でき、有意性の検定も共通となる。
・クロス表の縦横項目にはそれぞれ順番があるので、縦横が比例的な傾向のとき＋の相関係数となり、また、縦横が逆比例的な傾向のときーの相関係数となる。この点が属性相関係数は＋の値だけである事と異なる。
・順位相関＝０とはクロス表の縦横が無関連（独立）のときである。
・順位相関＝１とはクロス表の縦横が完全な関連の状態である。(付録６)
・一般の相関係数は、異常値があると、これに大きく左右されるが、Spearmanの順位相関は、順位を用いている為、異常値の影響が少ない。
・一般の相関係数は、密集した点に対する感度が低いが、Spearmanの順位相関は、密集点に対しても順位を用いている為、適当な感度を持つ。

2.2b　Spearmanの順位相関の有意性検定（図表2.2）

Spearmanの順位相関係数が０以外の有意な値を持つか否かの検定は、一般の相関係数と同じとなり、次の手順で行う。



図表2.2　Spearmanの順位相関の検定

①仮説として、「順序尺度から成る２変量は関連性が無く、従って相関係数は０である。」を立てる。
②対立仮説として、事前情報がある場合は、簡単のため、＋側を考え、「相関係数は＋である」とする。事前情報がない場合は、「相関係数は０ではない」とする。
③クロス表のT値を次式によって計算する。
　$T = Rs \times \sqrt{(N-2)/(1-Rs^2)}$　　ここで、RsはSperamanの相関係数
④自由度N－2のt分布において、上のT値に対する次の危険確率P値を求める。
　対立仮説が「相関係数は＋である」のとき、P値＝P(T以上)
　対立仮説が「相関係数は０ではない」のとき、P値＝P(±Tの外)
⑤危険確率P値≦0.05（許容される危険率５％）ならば、今回のデータから、相関係数は＋であると結論する。
⑥危険確率P値＞0.05ならば、データを増加して再検討する。または、仮説を採用して、今回のデータからは、ばらつきを考慮すれば、相関係数は０に近い、と結論する。

例題（所得階層と防衛意識）（図表2.3）

次の表は所得階層を３段階に区分して、各段階ごとに日本の防衛について、縮小、現状維持、拡大のどれが好ましいか、アンケートをとったという想定例である。調査に先立ち、所得が多いほど、生活に余裕があり、防衛も考えるだろうとの予想を立てたとする。

この例は社会的に微妙な問題を含んでいるが、このような時ほど、統計から言える範囲、言えない範囲を明確にする必要があるので、統計分析を適用する際の注意点を学ぶため、例題として取上げている。

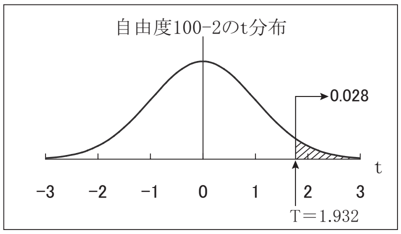
A表

	縮小	維持	拡大	計
低所得層	13	11	9	33
中所得層	13	8	12	33
高所得層	6	13	15	34
計	32	32	36	100

spearmanの相関係数　Rs=	0.192
相関有意の危険率　P＝	0.028
検定用t変数の値　T=	1.932

B表

X　平均順位	Y 16.5	48.5	82.5	計
17	13	11	9	33
80	13	8	12	33
83.5	6	13	15	34
計	32	32	36	100



図表2.3　所得階層と防衛についての意見：Spearmanの順位相関

・Ａ表は原因側と考えられる所得階層の各段階について調査人数をほぼ同等としている。Ｂ表では縦横項目に対する平均順位を併記した。専用ソフトによる結果は、
Spearmanの順位相関係数Rs＝0.192
クロス表のT値＝1.932
対立仮説「相関係数が＋の値」を採用するときの危険確率P値＝0.028

・P値は許容危険確率0.05（５％）より小さいので、このクロス表の数値からは、Spearmanの相関係数は正であると結論できる。すなわち、「このデータからは、所得が多い階層ほど、日本の防衛に積極的な意識を持つ傾向が見られた。また、比較データが無いので、この傾向の強弱は言えない。」

・ここでは相関係数の信頼区間については複雑となるので学ばないが、上の相関係数の計算値は与えられたデータからの平均的推定値である。微妙な性格を持つ問題に対しては本来、相関係数は確率的にばらつきを持つ事、すなわち推定幅を持つ事の理解が必要である。

2.3　一般の（ピアソンの）相関係数

一般に相関係数と呼ばれているのが、ここで取上げるピアソンの積率相関係数である。クロス表の縦横変数は、属性相関、順位相関の時と異なり数値データとなる。一方、各セル内は従来と同じ、該当データ個数となる。

ピアソンの相関係数は、Spearmanの相関係数と同じ計算式であり、順位変数の代わり、数値を次式のように用いる。また、有意性の検定はSpearmanの順位相関と全く同様なので、ここでは省略する。

クロス表にピアソンの相関係数を適用するときの注意点は、例で示すように、元は連続的な変数を、表の構成のため、区分しなければならない。すると、するとクロス表で扱う変数値は区分の代表値となることである。

ピアソンの順位相関係数

R＝縦横の共分散／（縦の標準偏差×横の標準偏差）

以下においてX、Yはクロス表の縦および横の変数である。

縦横の共分散＝Σ{N_{xy}×（X−AV_x）×（Y−AV_y）}／（N−1）

縦の標準偏差^2＝Σ{Nx×（X−AV x）^{2}}／（N−1）

横の標準偏差^2＝Σ{Ny×（Y−AV y）^{2}}／（N−1）

N_{xy}：　各セル内のデータ個数

N_x：　各セル内のデータ個数を横方向に合計

N_y：　各セル内のデータ個数を縦方向に合計

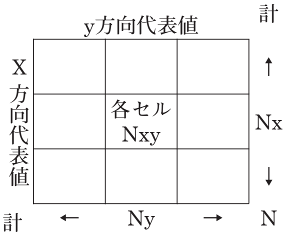
X：　縦方向変数の各代表値

Y：　横方向変数の各代表値

AVx：全Xの平均値

AVy：全Yの平均値

N：　セル内の全個数



縦横項目が数値の３×３クロス表

クロス表におけるピアソンの相関係数は、次の特徴を持っている。

・Spearmanの順位相関係数と計算式、有意性の検定が共通となる。

・クロス表の縦横が比例的な傾向のとき＋の相関係数となり、また、縦横が逆比例的な傾向のとき−の相関係数となる。

・相関係数＝０とはクロス表の縦横が無関連（独立）のときである。

・相関係数＝１とはクロス表の縦横が完全な関連の状態である。（付録６）

・データに飛び離れた値があると、相関係数はこれに大きく左右される。

・逆に密集した点に相関係数は左右されにくい。

例題（数学と英語の得点）（図表2.4）

次の表は、数学と英語の点数についての関連性を知るため、高校生100人の得点データを適切な方法で選んだとの想定である。従来の経験から２科目間に関連はあっても、弱い事が予想されるので、相関係数の有意性検定に重点をおきたい。

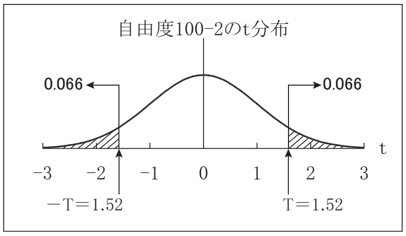
A表

	英語			
数学	0−40	41-70	71−100	
0−40	**13**	**12**	**8**	33
41−70	**10**	**14**	**9**	33
71−100	**9**	**11**	**14**	34
	32	37	31	100

B表

	Y			
X	20.0	55.0	85.0	
20	**13**	**12**	**8**	33
55	**10**	**14**	**9**	33
85	**9**	**11**	**14**	34
	32	37	31	100

ピアソンの相関係数	R＝	0.152
相関有意とする危険率	P＝	0.1320
検定用t変数の値	T＝	1.52



図表2.4　数学と英語の得点の相関

この例は２変数間の関連性を調べるのが目的で、因果関係は一般に考えられない例である。

A表は調査から得られたデータで、点数は範囲として示している。一方B表は点数を、代表値によって示し、相関係数が計算できる形としている。分析過程を示すと、

・仮説「数学と英語との相関係数は０である」を立てる。
　対立仮説「相関係数は０でない」を立てる
・専用ソフトより、ピアソンの相関係数R＝0.152
　クロス表のT値＝$0.152\times\sqrt{\{(100-2)/(1-0.152^2)\}}=1.52$

・対立仮説を採用したとき、危険確率P値＝P(±Tより外)＝２×0.066＝0.132
・P値は許容危険確率0.05（５％）より大きいので、対立仮説は採用できない。
　よって、数学と英語の得点間に関連性は見られないといえる。

・クロス表のP値は許容危険確率から大きく隔たっているので、このまま仮説を採用する。
　すなわち、データを増して対立仮説を再検討することは考えない。
・結論の統計的表現は、仮説「２変数間は関連性が無く、従って相関係数は０である」を受け入れる。よって、この相関係数値は、ばらつきを考慮すると０に近いと言える。

2.4 クロス表のまとめ

これまでクロス表の縦横変量が名義尺度、順序尺度、間隔尺度の各場合に分けて、相関係数と関連の有意性とを取上げてきたが、図表2.5に、これらのまとめを示す。

図表2.5 クロス表分析のまとめ

	名義尺度 （質的変数）	順序尺度 （順序変数）	間隔尺度 （数値変数）
検　定	←…… 縦横変数の関連性は有意？カイ自乗検定 ……→ ←…… 相関係数は0でないことの検定 ……→		
定義式	Cramerの属性相関 √（カイ自乗値／データ数／t）	Spearmanの順位相関 縦横の共分散／縦の標準偏差／横の標準偏差	ピアソンの積率相関
相関係数の性質	←…… 表の縦横が無関連なとき：相関係数＝0 ……→ ←…… 対角要素のみが値を持つとき：相関係数＝1 ……→ ←…… 表の縦横が等しくないとき：相関係数の最大値＜1 ……→		
正　負	常に相関係数＞0	←…… 表の縦横が比例的：相関係数＞0 ……→ ←…… 表の縦横が逆比例的：相関係数＜0 ……→	
散布図		小 ←…… 離れた点の影響 ……→ 大 大 ←…… 密集した点の影響 ……→ 小	

おわりに

ここにまとめたのは比較的最近の講義のうち、半期分と少しであるが、講義内容を記すだけでは、片手落ちのような気がする。講義に対する反応など、教室において気付いた点にも触れなければならないだろう。本稿の趣旨から外れない範囲で、これについて簡単に触れ、まとめとしたい。

筆者が学生の頃、始めて統計学に接した時の印象を思い出してみると、それまで中学、高校と習ってきた数学とは、かなり感覚の違う分野に思われた。おそらく、多くの学生にとって統計は、数学の１単元に現在は取り入れられているにしても、私と同様の印象を持っているものと推測している。従来の数学との感覚の違いの理由は、多分、扱う数値が一通りに定まらないことと、これによる結論も幅を持つ事にあるのだろう。

話は飛ぶようだが、他に担当の数学教育の試験などで、割り切れない数値が答となるような問題を出すと、クレームが良く出る。ましてデータがばらつく問題を扱うときの違和感に対してもっと早く教える側として気付くべきだったと思う。

統計学固有の「沢山の数で勝負」的な感覚に慣れてもらうため、授業の出発点として、各自が行うダーツ実験や、身の回りの例題作りを従来から行っているが、２集団の差の検定に至ると、これまで試みてきた工夫ではまだ不十分であることを感じる。これに対して

今考えられる事は、テーマ毎の統計モデルに合うよう、サイコロなどによる実験を考え、全員の結果を集めれば理論が正しい事がわかるような試みであろう。

クラスの人数が少ない時など、年によっては気楽な質問がいくつかのグループから出るようになり、用語などの不明な点から始まって、どのように誤解しやすいかについて伝わってくるようになって来た。時にはその場で、例えなどによる簡単な回答が、思いつかないような質問に接する事もあり、授業が参加型になるときもある。このような雰囲気を保つ為、皆が教室の前方に着席する事を薦め、各人の表情から理解度がうかがえるようにしているが、これも小規模大学ならではの、そして素朴な学生気質ならではの利点と感じている。最終的には、学生と教員との共同によって作り上げていく授業を目指したいと願っている。

参考文献

（１）安田三郎、他 「社会統計学」 丸善株式会社　1977
（２）杉山高一、牛沢賢二 「パソコンによる統計解析」朝倉書店1989
（３）脇本和昌、他 「パソコン統計解析ハンドブック」共立出版1995
（４）岩原信九朗 「教育と心理のための推計学」日本文化科学社1995
（５）東京大学教養学部統計学教室編 「基礎統計学Ⅰ　統計学入門」東京大学出版会1998
（６）竹内啓、他 「統計学辞典」東洋経済新報社2000

付録

１）例１の母集団は、県内の来春高校受験生の全員となるが、統一テストを受けた生徒は、母集団にほとんど等しいと考えられる。よってここでは、平均63点、標準偏差12点というのは、母集団についての推定値というより、真値となるようなモデルである。

２）例２の母集団は、20代の人全員となるので、この例における調査結果は母集団に対する推定値となる。調査に必要な人数を無作為に抽出するなど、適切に行われていれば推定値は正しいと見做せるが、この種の調査は、具体的な方法などに多くの課題を含むと考えられる。また、全数調査でないときの推定値は、ある幅を持つ。

３）例３の母集団は、携帯電話を利用している中学生全員となる。常時、通話料金を把握している電話会社にとって、他社の分は不明なので、全数調査とはならないが、調査結果は、母集団の正確な推定値となろう。

４）サンプル平均AV1、AV2はサンプルの度に確率的に変化する。そこで新たに｜AV1－AV2｜を変数としたときのばらつきを考えると（$SD_1^2 - SD_2^2$）／Nとなり、この$\sqrt{\ }$を合成標準偏差SDとしている。合成標準偏差SDが計算できる根拠は２集団に平均の差は無いという仮説、および母集団のばらつきは同一、との仮定より、両母集団を合わせて考える事ができるからである。

５）クロス表の縦横の個数が同一のとき、Cramerの相関係数は最大値＝１を取り、この状態を完全関連と呼んでいる。一方、クロス表の縦横の個数が異なるときは、Cramerの相関係数は最大値＜１となる。また、属性相関は一般に、Cramerの係数に限らず、その中において相対的大小は論じられるが、異なった係数を用いて比較するときは、P値などを用いて慎重に行う必要がある。

６）Cramerの相関係数と同様に、クロス表の縦横が同一区分のとき、Spearmanの相関係数およびピアソンの相関係数ともに最大値＝１となる。また、クロス表の縦横数が同じでないときは、Cramerの係数と同様に、最大値＜１となる。

A Study in Lectures on Educational Statistics (first report)

Tsutomu Doi

This report is one part of what the author has produced for the educational statistics classes since 1997. In this paper the focus is on four themes.

1. To generate the statistical distribution, each student carries out “darts” experiment in the classroom. After the experiment students learn some major theorems of the distribution.
2. Students are expected to make some application examples utilizing the statistical distribution and also expected to make some examples on the characteristics between two different population gropes.
3. The author has introduced a new topic on the study of cause-and-effect relation between two variables forming cross tables which are generated from questionnaire data base. It is emphasized that the variables are classified into numerical, ordinal and categorical types.
4. The modeling and calculation procedure is supported by Excel software which the author prepared for class use. This allow students to understand how to apply theory to data.