

# 社会システム分析のための統合化プログラム 4

## 基本統計

福井正康

福山平成大学経営学部経営情報学科

### 概要

我々は主として教育での利用を目的に、社会システム分析に用いられる様々な手法を統一的に扱うプログラムを作成しているが、今回は基本的な集計・検定・推定に関する統計手法をシステムに加えた。この論文では、用いられた分析手法と、それらの関係を図的に理解できるように作られた分析選択メニューについて詳述する。

### キーワード

統計，検定，統計的推定，社会システム分析，ソフトウェア，統合化プログラム

# 1章 はじめに

これまで社会システム分析で利用される数理的手法を統合的に扱うプログラムを MS-Windows 上の Visual Basic によって開発してきたが<sup>1-3)</sup>、この論文では統計処理のうち基本的な、集計、検定、推定等について、用いた理論とプログラムの構成をまとめて説明する。

統計処理ソフトウェアは、様々な機関で、人力と時間をかけて、数え切れないほど多く作成されており、個人が作るものにはおのずと限界がある。しかし、統合的な教育プログラムを作るという立場からは避けて通れない道であり、その際ある種の独自性を打ち出す必要もある。

統計処理プログラムは一般に個々の分析プログラムの集合体となっており、ユーザーは必要に応じてそれらを選択して使い分ける。しかし、統計に不慣れな初心者にとってはどの分析をどのように利用するか、その判断こそが最も難しい。しかし、自分が行おうとする分析の位置付けが明確に示され、その指針がプログラム中にあれば、判断の手助けとなり、安心感を持って分析が実行できるに違いない。特に統計学の講義を受講している学生にとっては、このガイドラインが必要であろう。

分析の位置付けを明らかにするという考え方は主に検定手続きの中で実現されている。検定の体系（異論のある方もおられるかも知れないが）を図式化したメニューをダイアログボックスとして示し、その中から自分の利用する分析手法を選択する。この考え方は特に目新しいものではないが、必ずや学習の手助けになるものと信じる。

このシステム中で利用できる統計処理手法は、データの集計、検定値と確率間の変換、質的指標の検定、量的指標の検定、標本数の決定、区間推定、相関係数と回帰分析と、大きく7つに分けられている。まず、データの集計では、具体的な調査データから分割表を作成したり、基本統計量を計算する。また、統計学の基礎知識として、分布と確率の概念について理解を助けるために、検定値と上側確率の相互変換の機能を持ったメニューも用意されている。統計的検定では、母集団と標本の比較及び、2標本間の差の検定が中心になっている。ここでは広範囲に利用できるように、最低限のノンパラメトリック検定も含んでいる。入力用のデータについては、加工されていない調査データや他で計算された基本統計量を元にするを想定し、実用的な利用から講義等での計算チェックまで、幅広い需要に対応させる。また、アンケート調査等で標本数を決める必要性も考えて、標本を母集団と比較するために必要な標本数の決定についてのメニューも用意している。推定は、母比率及び、データの正規性を仮定した場合の母平均と母分散の推定を含んでいる。最後に、2変量の関係を明らかにするために、相関係数の検定と線形回帰分析が加えられている。

欠損値データの処理方法、有意水準の指定と片側・両側検定の区別、エディターからの変数の選択については、共通の設定項目としてコマンドボタンにより各分析から簡単に設定できるようになっている。これらには適当なデフォルト値が与えられ、分析に不都合が生じないよう

になっている。

## 2章 システム構成

### 2.1 設定

統計分析について説明を始める前に、欠損値の処理、有意水準の設定、エディター上の変数の選択方法に関する設定事項について述べる。実行画面は図1で与えられるが、このメニューは各分析から共通に呼び出され、この中で指定された設定はプログラムの実行中値が保持される。

欠損値の除去方法は、選択された変数についてのレコード単位の除去、各データ毎の個別の除去、統計手法に応じた自動選択がある。有意水準の設定については、片側検定、両側検定、検定手法に応じて標準的なものを選択する自動選択がある。例えば、 $t$ 検定とF検定は片側検定であり、 $t$ 検定その他については両側検定である。その数値は、パーセント表示で入力するが、デフォルトは5%になっている。もちろん次節の集計等のように有意水準に無関係なものについて、この値は無視される。

変数選択によって、エディター上のデータから利用される変数が選ばれるが、左側のリストで変数名を選択することによって、それが右側のテキストボックスにカンマ区切りで現れる。初心者のために、分析毎にデフォルトの選択が用意されているので、必要なデータだけをエディターで作成すると、設定画面を開くことなくボタン1つで統計処理が実行されるようになっている。

### 2.2 データの集計

データの集計は、質的指標と量的指標に分けて処理される。具体的な実行画面は図2で与えられる。

質的指標では、項目ごとにデータ数を集計し、分割表が作られる。たとえば、表1のような1次元データは、その下に見られるように1次元分割表として集計される。また、表2のような2次元データは、2次元分割表として集計される。もちろん、項目名が文字列で表わされている場合も同様である。行と列の関係は設定の変数選択の順番で決まる。現在、分割は2次元分割表までである。これらの分割表は、



図1 初期設定画面



図2 データの集計

質的指標の検定のところでも作成することができる。

表1 1次元分割表

Field1: 1, 2, 1, 1, 2, 2, 3, 1, 2, 1,  
3, 2, 1, 1, 1, 2, 3, 1, 3, 1

	度数
1	10
2	6
3	4
計	20

表2 2次元分割表

Field1: a, a, b, b, a, b, a, a, b, a, b, b, a, b, b, a, b, a, b, a  
Field2: 1, 2, 1, 1, 2, 2, 3, 1, 2, 1, 3, 2, 1, 1, 1, 2, 3, 1, 3, 1

	1	2	3	計
a	6	3	1	10
b	4	3	3	10
計	10	6	4	20

量的指標の場合は、基本統計量と度数分布表・ヒストグラムを表示する。基本統計量についてその項目と定義は表3で与えられる。不足する部分については随時追加する予定である。

量的データの分布型を見るために、度数分布表とヒストグラムを表示させる。度数分布表には、度数・相対度数・累積度数・累積相対度数が含まれる。設定は初期値と分割幅を指定するようになっている。

ここではさらに、入力データのレコードをある指標によって分割しながら、基本統計量や度数分布表等を求める機能を追加する必要がある。

基本統計量については殆どの検定メニューから、度数分布表とヒストグラムについては正規性の検定のメニューから見る事ができる。

表3 基本統計量

データ数	$n$
平均値	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$
中間値	$Me$
最大値	$\max\{x_i\}$
最小値	$\min\{x_i\}$
範囲	$\max\{x_i\} - \min\{x_i\}$
分散	$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$
不偏分散	$u^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$
標準偏差	$s$ または $u$
歪度	$a_3 = \frac{1}{n} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{s} \right)^3$
尖度	$a_4 = \frac{1}{n} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{s} \right)^4$

### 2.3 検定値と確率

基本的な分布関数について、教育用に検定値から上側確率及び、上側確率から検定値を求める必要があり、簡単な計算メニューを加える。具体的な実行画面は、図3で与えられる。ここではパラメトリックな検定に利用される、標準正規分布、<sup>2</sup>分布、F分布、t分布について、結果が求められる。具体的な計算には、文献4)の計算式を利用した、文献1)の補遺4に与えられる確率関数を用いた。2項分布、多項分布、Poisson分

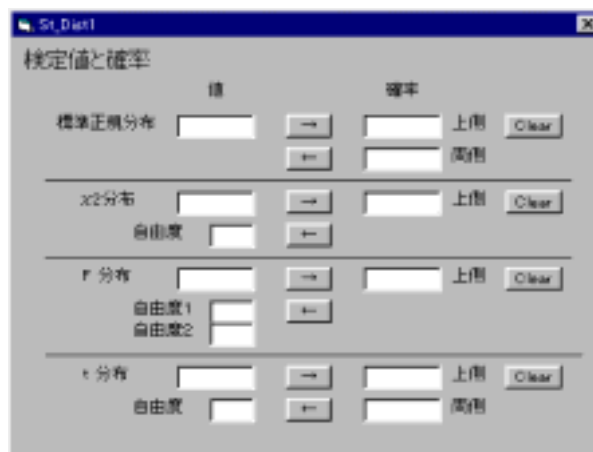


図3 検定値と確率画面

2項分布、多項分布、Poisson 分布、指数分布等、その他の分布についても確率値が簡単に求められるメニューを作成しなければならない。

## 2.4 質的指標の検定

質的指標の検定手順については、  
図 4 の分類を用いた<sup>3)</sup>。データ数の少ない場合など、この考え方が利用できないこともあるが、その対応は今後の課題とする。

利用者に検定手法の位置付けを明確に認識させるために、分析を選択するメニューを一般的な統計ソフトで見られる羅列的なものとせず、図 4 の形式をそのままメニュー化した。

具体的な実行画面を図 5 に示す。図 5 の検定のコマンドボタンから具体的な分析メニューが呼び出される。利用する分布公式については、図 4 の検定手法に応じて表 4 のようにまとめられる。

適合度検定について、図 6 に実行画面を示す。一般に、質的指標の検定には 2 種類の検定用データが考えられる。1 つは調査票等から直接入力された表 1 や表 2 の形式をしたデータで、それを元に分割表の作成や検定が行われる。また既に分割表を作成している場合には、その分割表を利用して検定を実施することも考えられる。実際の調査等では前者の形式が多くなるであろうが、講義用としては後者の場合も必要である。

それゆえ、このプログラムでは質的指標の検定の際、どちらかのデータ形式を選択するように

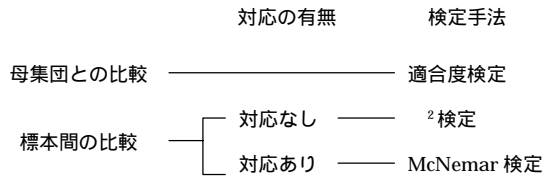


図 4 質的指標に関する検定手法の分類



図 5 質的指標の検定画面

表 4 質的指標の検定

適合度検定
標本数 $n$ , 事象 $i$ の出現回数 $n_i$ , 事象 $i$ の母比率 $p_i$
$\chi^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i} \sim \chi_{k-1}^2 \text{ 分布}$
$\chi^2$ 検定
標本数 $n$ , 要因 $i$ 事象 $j$ の出現回数 $n_{ij}$ , $n_{.j} = \sum_{i=1}^r n_{ij}$ , $n_{i.} = \sum_{j=1}^s n_{ij}$
$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - n_{i.}n_{.j}/n - 1/2)^2}{n} \sim \chi_{(r-1)(s-1)}^2 \text{ 分布}$
特に、 $r = k = 2$ のとき、 $\chi^2 = \frac{n( n_{11}n_{22} - n_{12}n_{21}  - n/2)^2}{n_{1.}n_{.2}n_{1.}n_{.2}} \sim \chi_1^2 \text{ 分布}$
McNemar 検定
群・対照群の要因の有無別数 ( 有有 $a$ , 有無 $b$ , 無有 $c$ , 無無 $d$ )
$\chi^2 = \frac{( b - c  - 1)^2}{b + c} \sim \chi_1^2 \text{ 分布}$

なっている。前者のデータの場合、分割表だけを作る場合もあると考えられるので、これらの検定メニューからも分割表が作れるようになっている。

実測値と比較する理論確率については、カンマ区切りで入力する。例えば、0.5, 0.3, 0.2 のような小数表示と 1/3, 1/3, 1/3 のような分数表示が可能である。メニューには注意書きを多く加え、分かり易さを高めている。

また、正規分布の検定値を用いた比率の検定は分割数が 2 の適合度検定と同等であるので、省略している。

結果表示には検定結果の数値表示の他に、初心者の学習用に、例えば「標本値と理論値とを比べて差があるといえない。」のような検定結果を言葉にした表現や、標本数に関する利用上の注意等を加える。

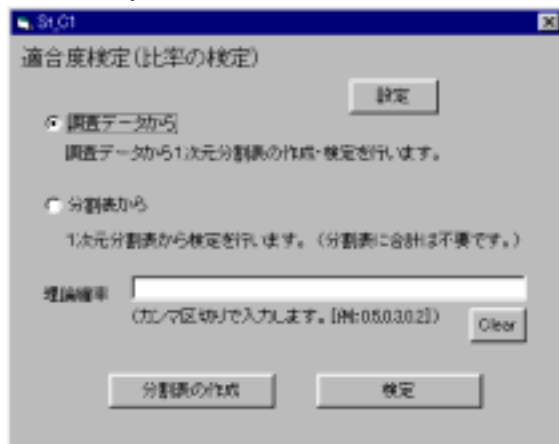


図 6 適合度検定画面

## 2.5 量的指標の検定

量的指標の場合には図 7 の分類法と検定手法を用いる<sup>5)</sup>。特に、ノンパラメトリック検定についての他の分析手法や、適用限界についてのさらに細かい分類は今後の課題とする。

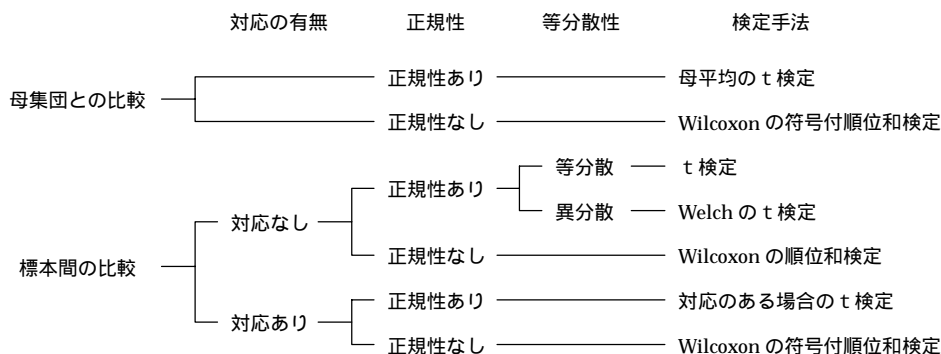


図 7 量的指標に関する検定の分類

質的指標と同様に、量的指標に関しても検定の位置付けを明確にするために、図 7 の様式を持った検定メニューが用意されている。その実行画面は図 8 で与えられる。このメニューでは、右端の検定手法だけでなく、分類項目である正規性の検定や等分散性の検定も選択できるようになっている。

ここでは検定手法を母集団との比較と標本間の比較とに分け、標本間の比較については、そ

これらの間の対応の有無によってさらに分類する。

量的指標の検定の基本性質は、パラメトリック検定とノンパラメトリック検定を分ける分布の正規性であるが、これらの見極めのために正規性の検定が必要である。そのために、ここでは目視的方法と数値的方法の2通りを用意する。

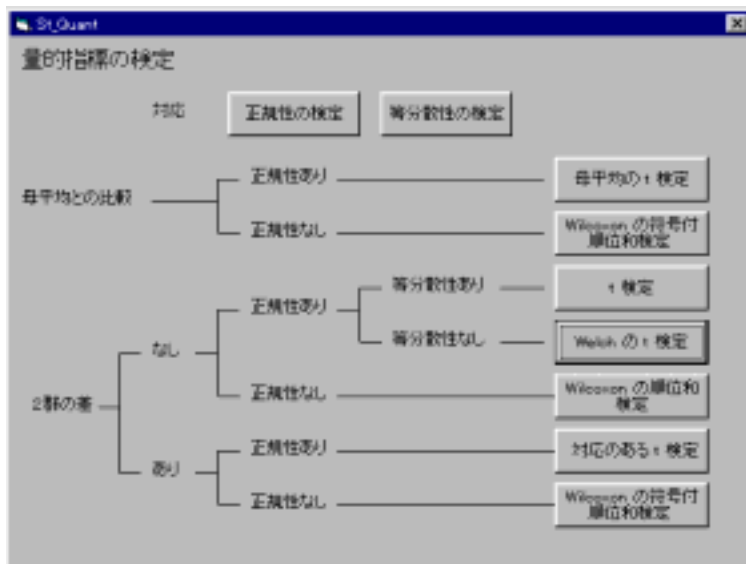


図 8 量的指標の検定画面

目視的方法としては、データ数が多い場合に使われる、度数分布表やヒストグラムから正規性を見る方法、またデータ数が少ない場合に利用される、正規確率紙による方法が用意されている。グラフは正規確率紙へのプロットに準じて、データの個数を  $n$ 、あるデータの順位を  $i$  としてその累積確率を  $i/(n+1)$  で与え、データの数値と、この累積確率から得られる標準正規分布の検定値とで分布図を描く。これに回帰直線を加え、直線状への並びを見易くする。

正規性の数値的な検定方法としては Kolmogorov-Smirnov 検定を利用するが、Shapiro-Wilk の W 統計量等の検定についても考察中である。

母集団との比較に関して、その手法は表 5 にまとめる。非正規性の場合の検定は分布の対称性を仮定して、Wilcoxon の符号付順位和検定を採用した<sup>6), 7)</sup>。またこの検定において、同順位の場合は順位平均を用いるが、同順位が多く含まれる場合の補正は今後の課題とする。

表 5 量的指標の検定 (母平均との比較)

<p>母平均の t 検定          標本数 <math>n</math> , 標本平均 <math>\bar{x}</math> , 不偏分散 <math>u^2</math> , 母平均 <math>\mu</math></p> $t = \frac{\sqrt{n} \bar{x} - \mu }{u} \sim t_{n-1} \text{ 分布}$ <p>Wilcoxon の符号付順位和検定          データ <math>x_i</math> , 中間値 <math>\mu</math> , <math>z_i = x_i - \mu</math>  <math> z_i </math> の昇順に 0 を除いて順位 <math>r_i</math> を付け、<math>z_i</math> の正負で 2 群に分類          各群の順位和 <math>R_r</math> , <math>R_s</math> の中で小さい方を選択 <math>R = \min(R_r, R_s)</math>          標本数が少ないとき (<math>z_i \neq 0</math> の例数 <math>&lt; 10</math>)          文献 5), 6) 等の数表の利用          標本数が多いとき (<math>z_i \neq 0</math> の例数 <math>\geq 10</math>)</p> $z = \frac{R - n(n+1)/4}{n(n+1)(2n+1)/24} \sim N(0,1) \text{ 分布}$
--

標本間の比較の場合は、対応のある場合とない場合とに分類する。対応がない場合、正規性の検定を行い、正規分布ならさらに等分散性を検定する必要がある。これらの分類による具体的な検定手法は表 6 にまとめる。正規性の認められない場合は Wilcoxon の順位和検定を用いる。

表 6 量的指標の検定 (対応のない 2 標本の比較)

<p>F 検定 (等分散性の検定)</p> <p>標本数 <math>n_1, n_2</math>, 不偏分散 <math>u_1^2, u_2^2</math> (<math>u_1^2 &gt; u_2^2</math>)</p> $F = \frac{u_1^2}{u_2^2} \sim F_{n_1-1, n_2-1} \text{ 分布}$ <p>(student の) t 検定</p> <p>標本数 <math>n_1, n_2</math>, 標本平均 <math>\bar{x}_1, \bar{x}_2</math>, 不偏分散 <math>u_1^2, u_2^2</math></p> $t = \frac{\sqrt{\frac{n_1 n_2}{n_1 + n_2}} \frac{ \bar{x}_1 - \bar{x}_2 }{\sqrt{\frac{(n_1 - 1)u_1^2 + (n_2 - 1)u_2^2}{n_1 + n_2 - 2}}}}{\sim t_{n_1 + n_2 - 2} \text{ 分布}}$ <p>Welch の t 検定</p> <p>標本数 <math>n_1, n_2</math>, 標本平均 <math>\bar{x}_1, \bar{x}_2</math>, 不偏分散 <math>u_1^2, u_2^2</math></p> <p>自由度 <math>d = \frac{1}{\frac{c^2}{n_1 - 1} + \frac{(1 - c)^2}{n_2 - 1}}</math>, <math>c = \frac{u_1^2/n_1}{u_1^2/n_1 + u_2^2/n_2}</math></p> $t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{u_1^2/n_1 + u_2^2/n_2}} \sim t_d \text{ 分布}$	<p>Wilcoxon の順位和検定</p> <p>標本数 <math>n_1, n_2</math> (<math>n_1 \leq n_2</math>), 標本 <math>x_i^1, x_j^2</math></p> <p>標本の昇順に順位 <math>r_i</math> を付け、標本数の少ない群の順位和を求める。</p> $W = \sum_{i=1}^{n_1} r_i$ <p>標本数が少ない場合 (<math>n_2 \leq 20</math>) 文献 5), 6) 等の数表を利用</p> <p>標本数が多い場合 (<math>n_2 &gt; 20</math>)</p> $Z = \frac{W - \frac{1}{2}n_1(n_1 + n_2 + 1)}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \sim N(0, 1) \text{ 分布}$
--	--

対応のない 2 標本の比較の場合、データの読み込み方法は 3 種類用意する。例として図 9 に正規性と等分散性を仮定した t 検定のメニュー画面を示す。調査データには通常、各レコードにデータを区分するための質的データと量的データが含まれている。この質的データを元にレコードを分類し、量的データについて検定を行うが、その際、元データから直接読み込む場合、既に群別に分けてあるデータを元にする場合、さらに群別データの基本統計量が求められている場合について検定が実施できるようにする。これは教育用には必須の機能であると考えられる。

対応のある場合の検定手法は表 7 にまとめる。この場合、調査データ自体、群別に分類されており、群別データを考える必要

がある場合の検定手法は表 7 にまとめる。この場合、調査データ自体、群別に分類されており、群別データを考える必要

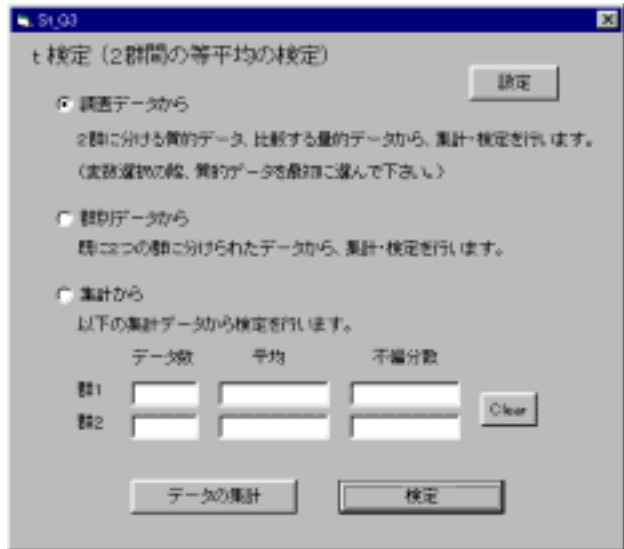


図 9 t 検定画面

表 7 量的指標の検定 (対応のある 2 標本の比較)

<p>対応がある場合の t 検定</p> <p>例数 <math>n</math>, 標本差 <math>z_i</math>, 平均 <math>\bar{z}</math>, 不偏分散 <math>u_z^2</math></p> $t = \frac{\sqrt{n}  \bar{z} }{u_z} \sim t_{n-1} \text{ 分布}$ <p>Wilcoxon の符号付き順位和検定</p> <p>標本差 <math>z_i</math> をもとに表 5 の手法を利用する。</p>
---



表 8 標本数の決定

はなく、入力は2種類でよい。検定は対応のある各レコード毎の標本差を利用するが、正規性の認められない場合もこれを用いて、表5で示したWilcoxonの符号付順位和検定を利用する。

## 2.6 標本数の決定

標本数の決定については、正規性が認められる場合に限定し、母比率の検定と母平均の検定のために必要なデータ数を求める。具体的な公式は表8にまとめる。但し、母平均を求める検定に必要な標本数は、数が多いものとして近似的に標準正規分布の検定統計値を利用している。ここに、 $\chi_1^2(\alpha)$ は自由度1の $\chi^2$ 分布の上側確率 $\alpha$ の検定統計値であり、 $Z(\alpha/2)$ は標準正規分布の上側確率 $\alpha/2$ の検定統計値である。

質的指標で分割数が3以上の場合や2群間の差の検定及び、正規性を持たない場合等の標本数の決定については今後の課題とする。図10に標本数の決定の画面を示すが、入力には母集団の統計量と、データを収集した場合の予想値とを用いる。標本数の決定に関しては、予想値によるところが大きいので、多くの検定手法への対応は特に重要であるとは考えない。

## 2.7 区間推定

区間推定についても正規性が認められる場合に限定する。求める推定値は、母比率、母平均、母分散とした。具体的な手法については、表9にまとめる。ここに、前節で説明した表式を除いて、 $t_{n-1}(\alpha/2)$ は自由度 $n-1$ の $t$ 分布の上側確率 $\alpha/2$ の検定統計値である。表式の簡単化のために、母比率と母平均については上限と下限を示すこととする。

入力は調査データからの入力と統計量からの入力と2種類持っておけばよい。正規性が認められない場合の推定については今後の課題とする。

母比率の検定用 母比率 $p$ , 標本比率 $\hat{p}$
$n = \frac{\chi_1^2(\alpha)p(1-p)}{(\hat{p}-p)^2}$
母平均の検定用(両側) 母平均 $\mu$ , 母分散 $\sigma^2$ , 標本平均 $\bar{x}$
$n = \frac{Z(\alpha/2)^2\sigma^2}{ \bar{x}-\mu ^2}$

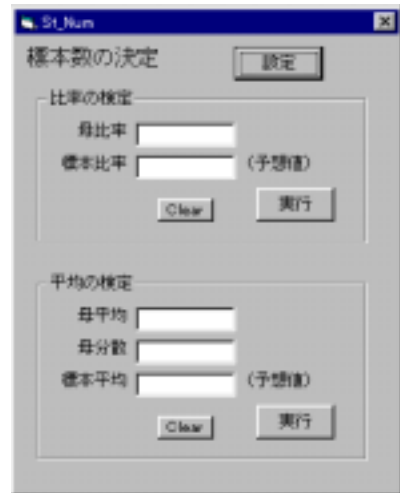


図 10 標本数の決定画面

表 9 区間推定

母比率の推定 標本数 $n$ , 標本比率 $\hat{p}$
$\hat{p} \pm Z(\alpha/2)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
母平均の推定 標本数 $n$ , 標本平均 $\bar{x}$ , 不偏分散 $u^2$
$\bar{x} \pm \frac{u}{\sqrt{n}}t_{n-1}(\alpha/2)$
母分散の推定 標本数 $n$ , 不偏分散 $u^2$ , 母平均 $\sigma^2$
$\frac{(n-1)u^2}{\chi_{n-1}^2(\alpha/2)} \leq \sigma^2 \leq \frac{(n-1)u^2}{\chi_{n-1}^2(1-\alpha/2)}$

## 2.8 相関係数と回帰分析

相関係数については、正規性が認められる場合の Pearson の相関係数及び、正規性が認められない場合の Spearman の順位相関係数について求めており、無相関か否かの検定を行っている。また、回帰分析については、回帰式と重相関係数、及び寄与率について求め、回帰係数の有効性について、残差の正規性を仮定して検定を行っている。また、結果表示には回帰直線も含めた分布図も利用する。具体的な公式については表 10 にまとめる。説明変数は 1 つだけに限り、複数の場合は重回帰分析として多変量解析に譲る。

表 10 相関係数と回帰分析

標本数 $n$ , 相関係数 $r$	$t = \frac{ r  \sqrt{n-2}}{\sqrt{1-r^2}} \sim t_{n-2} \text{ 分布}$
Spearman の相関係数の検定	
標本数 $n$ , 群ごとの順位による順位相関係数 $r_s$	$t = \frac{ r_s  \sqrt{n-2}}{\sqrt{1-r_s^2}} \sim t_{n-2} \text{ 分布}$
回帰分析	
標本平均 $\bar{x}, \bar{y}$ , 不偏分散 $u_x^2, u_y^2$ , 相関係数 $r$	$y = ax + b , \quad a = r \frac{u_y}{u_x} , \quad b = \bar{y} - r \frac{u_y}{u_x} \bar{x}$
重相関係数 $R$	実測値 $y_i$ と予測値の相関係数
寄与率	$R^2$

回帰分析の検定については、表中では表しにくいので、ここで簡単にふれておく。目的変数を  $y$ 、説明変数を  $x$  とし、これらに、関係式  $y = ax + b + \varepsilon$  があると仮定する。ここに予測式は  $Y = ax + b$  であり、残差は  $\varepsilon \sim N(0, \sigma^2)$  分布とする。

回帰係数の有効性の検定は、データ数  $n$ 、残差変動  $EV = \sum_{i=1}^n (y_i - Y_i)^2$ 、説明変数の不偏分散  $u_x^2$  として、以下の関係を用いる。

$$t_a = \frac{a}{\sqrt{\frac{EV}{n-2} / (n-1)u_x^2}} \sim t_{n-2} \text{ 分布} \quad t_b = \frac{b}{\sqrt{\frac{EV}{n-2} \left( \frac{1}{n} + \frac{\bar{x}^2}{(n-1)u_x^2} \right)}} \sim t_{n-2} \text{ 分布}$$

単回帰分析の場合に前者の検定は、残差変動に対する回帰変動の有効性を検定する、回帰式の有効性の検定と一致する。

## 3章 考察と今後の展望

これまで、基本統計として社会系学部 of 1 年間の講義として無理のない程度の検定や推定を考えてきたが、実用上には多くの課題が残っている。特にノンパラメトリック統計の分野では、ここで取り上げたもの以外に多くの手法が確立されているが、初期の構想からは除外した。これらの手法を網羅して行くことはこのプログラムの本意ではなく、本格的な統計アプリケーションにまかせるべきであろう。ここでは社会システム分析手法について学習中、もしくは少し理解できた程度の利用者に対象を絞り、浅くではあるが広い分野に適用できるプログラムにしなければ、他の市販アプリケーションに太刀打ちできない。そのためには分かり易さを優先させ、多少の厳密性の犠牲には目をつむらざるを得ない。ただし、利用者の誤用を避けるため

に、理論の適用範囲を明確にする姿勢だけは保ちつづける必要がある。

統計処理を行う場合、分析プログラムに加えてエディターの機能強化が求められる。その例として例えば、入力された数値を利用しての計算機能がある。データの中にはそのままでは正規性を示さないがある変換式を用いると正規分布するものもあるし、理論式の値を求め、実測値との差を検討する場合もある。計算機能はこれらのデータ作りに利用されるため、エディターの機能としては必要不可欠である。また、あるキーを用いたレコードのソート機能も重要である。群別データ作りは統計処理としてメニューから選択するだけでなく、この機能により手軽に実現される。また、ある条件に見合うレコードを抽出する場合も、ソート機能を用いてある程度実現できる。その他にも多くの追加すべき機能があるが、これらは特に重要である。

統計関数を多く含んだ Microsoft Excel 等の表計算ソフトウェアは、統計学の学習に効果的である。自分が計算すべき統計量が認識でき、手軽に計算できるところは、高価な統計ソフトウェアよりむしろ教育的ともいえる。また例えば Excel には分析ツールとして、ある程度の統計処理のできるアドイン機能も含まれている。さらにこれらの表計算ソフトウェアは最近では殆どのパソコンに予めインストールされており、誰でも手軽に利用できる。我々のプログラムは基本的に複数枚のシートからなり、表計算ソフトウェアに似た構造をしている。それゆえ、特に統計分野などは Excel 上のアドインプログラムにすることに適しており、現にいくつかのアプリケーションも販売されている。ただ、使い易さと今後の様々な分析に対応させる柔軟さを考えた場合、エディターの機能強化を図り独立のアプリケーションにするか、それとも Excel に頼るべきか、難しい選択である。

## 参考文献

- 1) 福井正康, 田口賢士, 社会システム分析のための統合化プログラム, 福山平成大学経営情報研究, 3号, 109-127, 1998.
- 2) 福井正康, 田口賢士, 社会システム分析のための統合化プログラム 2 - 産業関連分析・KSIM・AHP -, 福山平成大学経営情報研究, 3号, 129-144, 1998.
- 3) 福井正康・増川純一, 社会システム分析のための統合化プログラム 3 - 線形計画法・待ち行列シミュレーション -, 福山平成大学経営情報研究, 4号, 99-115, 1999.
- 4) 脇本和昌・垂水共之・田中豊編, パソコン統計解析ハンドブック 基礎統計編, 共立出版, 1984.
- 5) 丹後俊郎・古川俊之監修, 新版 医学への統計学, 朝倉書店, 1993.
- 6) 柳川堯, 新統計学シリーズ 9 ノンパラメトリック法, 培風館, 1982.
- 7) 白旗慎吾編, パソコン統計解析ハンドブック ノンパラメトリック編, 共立出版, 1987.

# Multi-purpose Program for Social System Analysis 4

- Elementary Statistics -

Masayasu FUKUI

Department of Management Information, Faculty of Management,  
Fukuyama Heisei University

## **Abstract**

The author is constructing a unified program on social system analysis for the purpose of education. In this time, programs on the elementary statistical analysis are added to our system. The aim of this paper is explaining the methods of statistical test adopted in our program and concept of graphical menu illustrating the relation between them.

## **Keywords**

statistics, statistical test, statistical estimation, social system analysis, software, unified program