

第9章 実験データの整理

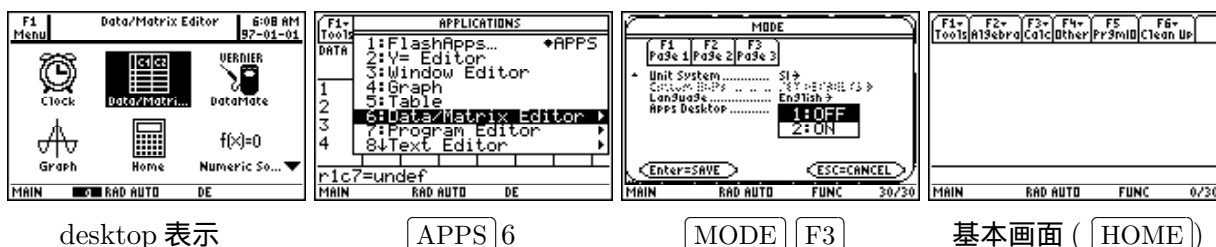
数ナビ (TI-89Titanium) には、エクセル (Excel) と同等の表計算機能が含まれています。その機能を利用すると、種々の表計算のみならず、そのデータをグラフ表示したり、最小自乗法などを利用した種々の統計解析を行うことができます。

9.1 Data/Matrix Editor の起動

Data/Matrix Editor の起動 / ファイル選択画面

データ処理を行うには、「Data/Matrix Editor」を起動させる必要があります。電源を入れた直後の画面に「Data/Matrix Editor」のアイコンがあるので、それを起動させてください。アイコンを「表示しない」設定にしている場合は、**[APPS]**を押すとメニューが出るので「6: Data/Matrix Editor」を選択します。

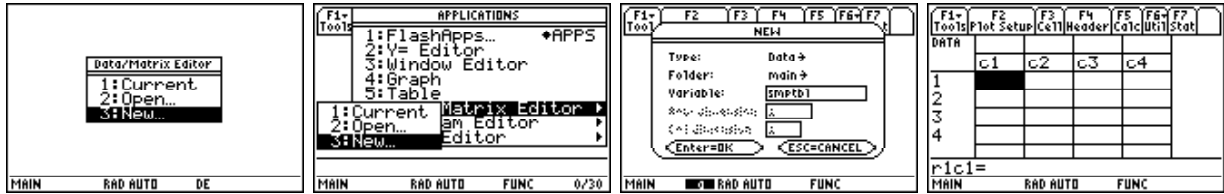
なお、初期画面にアイコンを表示させるかどうかは、**[MODE]** **[F3]** を押して一番下に表示される「Apps Desktop」の箇所で切り替えることができます。デフォルトでは「ON」に設定されているので、「OFF」に変更すると、次に電源を入れると最初から基本画面 (**[HOME]**) が表示されます。



ファイル選択画面

「Data/Matrix Editor」を起動すると、最初に既存ファイルをオープンするのか、それとも新規にファイルをオープンするのかを問われます。そのファイルの選択画面では、最後にオープンしたファイルを扱うときは「1: Current」を、すでに保存済みのファイルを扱うときは「2: Open」を、新規にファイルを作成するときは「3: New」を選択します。

「1: Current」を選択すると、最後にオープンしたファイルが開きます。「2: Open」を選択すると、既存ファイルのリストが表示されるので、その中から希望するファイルを選択します。「3: New」を選択すると、最初にファイル名を「Variable」の箇所に登録する必要があります。適当なファイル名 (例えば、「smptbl」) を入れて **[ENTER]** を押すと、空白の表が表示されます。メモリーの容量にもよりますが、最大で縦 999 行、横 99 列まで扱うことができます。



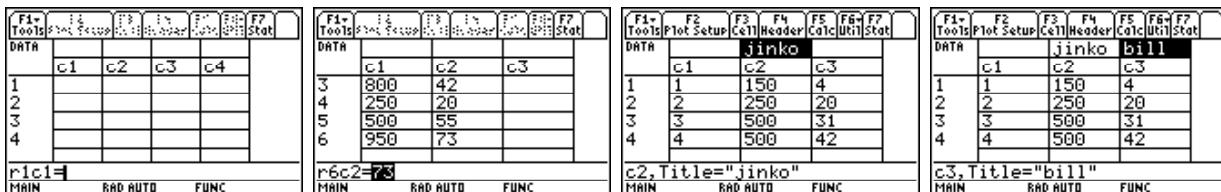
desktop からの open [APPS]6 での open ファイル名の入力 空の表

9.2 データの入力

表の個々のセルには、数値・文字・関数などを入力することができます。エクセルと違って数式処理ができるので、たとえば1列目にいろいろな関数を定義して、2列目ではそれを微分したりすることが可能です。ここでは、エクセルと同じように、数値を入力する場合について説明します。

まず、入力したいセルに で移動して [ENTER] を押すと、カーソルが入力行(一番下の行)に移動するので、そこで数値を入れて [ENTER] を押すと、指定したセルに数値が入力されます。例えば、マニュアル(248頁)にある右のような表を作成してみましょう。これは、マニュアルによれば、7つの都市の人口(千人)と、その都市の12階以上のビルの数のデータです。C1列に人口を、C2列にはビルの数を入れます。表の一番上の空白の欄は、その列が何のデータであるかが分かるような名前(タイトル)をつける欄です。C1列には「jinkou」、C2列には「bill」を入れておきます。

| | 人口 | ビル数 |
|---|-----|-----|
| 1 | 150 | 4 |
| 2 | 500 | 31 |
| 3 | 800 | 42 |
| 4 | 250 | 20 |
| 5 | 500 | 55 |
| 6 | 950 | 73 |



カーソルを入力行に 値の入力 1列目のタイトル 2列目のタイトル

9.3 平均値などの計算

この表を作成しておくで、それをもとにいろいろな統計上の計算をすることができ、列ごとの平均などは簡単に計算できます。[F5] を押すと、そのメニューが表示されます。

「Calculation Type」の箇所 を押すと、このメニューから計算できる統計計算のいろいろな種別が表示されます。1つの列の平均を求めるには「1: One Variable」を選択し、平均を求める列を指定します。列は左から順に、C1, C2, ... となります。C1列を指定して [ENTER] を2回押すと、C1列に関する多くの統計量が表示されます。 を押すことで、残りのものが表示されます。四分位数、メジアン、標準偏差などについては、5年の確率統計で学びます。

「2: Two Variable」を指定すれば、指定した2つの列に対して同様の統計量が表示されます。

| | | | |
|------------|--------------|----------|---------------|
| \bar{x} | 平均 | $\sum x$ | 列の合計 |
| $\sum x^2$ | 2乗 x^2 の合計 | sx | 標本標準偏差 |
| σx | 母集団標準偏差 | nStat | データの数 |
| minX | 最小値 | q_1 | 第1四分位数 (25%点) |
| medStat | メジアン (50%点) | q_3 | 第3四分位数 (75%点) |
| maxX | 最大値 | | |



F5 のメニュー

1変数を指定

C1列を指定

結果の表示

9.4 グラフ表示

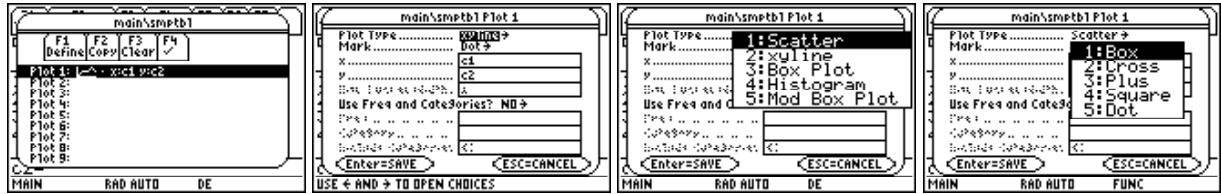
入力した2つのデータは、それぞれが無関係な2種類の値ではなく、1つの行の数値はそれぞれが対応した数値です。たとえば、人口150(千人)の都市では12階以上のビルが4つあり、人口500(千人)の都市では31のビルがあることが示されています。数ナビのグラフ機能を利用すれば、これらの間の対応関係を簡単にグラフ化することができます。その手順は、下記の通りです。

- (1) F2の「Plot Setup」を押すと、どのようなグラフ表示にするかを指定する画面が現れます。
- (2) Plot1からPlot9まで、9通りのグラフを定義できます。まず、どの箇所に定義するかを指定します。デフォルトでは、Plot1ではx軸はC1列、y軸はC2列になっています。ここではそのままが良いので、Plot1が黒く反転していることを確認の上、その箇所にさらに詳細を定義するためF1「Define」を押します。
- (3) すると、どのようにグラフ表示するかを指定する画面に切り替わります。一番上の「Plot Type」の箇所で、どのようなグラフにするかを指定します。例えば、「1: Scatter」を選択します。

| | |
|--------------|--------------------------------|
| 1: Scatter | 指定した2つの列のデータ(x,y)を点(ドット)で表示する。 |
| 2: xyline | 点と点の間を線(line)で結んで表示する。 |
| 3: Box Plot | 統計で学ぶ、「箱ひげ図」を表示する。 |
| 4: Histogram | 度数分布表の棒グラフを表示する。 |

- (4) 「Mark」の箇所では、点(x,y)をどのような形で表示するかを指定することができます。ここでは、例えば、「1: Box」を指定します。

| | | | |
|---------|------------|-----------|------------|
| 1: Box | 点に 印を置く | 2: cross | 点に × 印を置く。 |
| 3: Plus | 点に + 印を置く。 | 4: Square | 点に 印を置く。 |
| 5: Dot | 点に ・ 印を置く。 | | |



[F2] Plot Setup [F1] Define 1: Scatter 1: Box

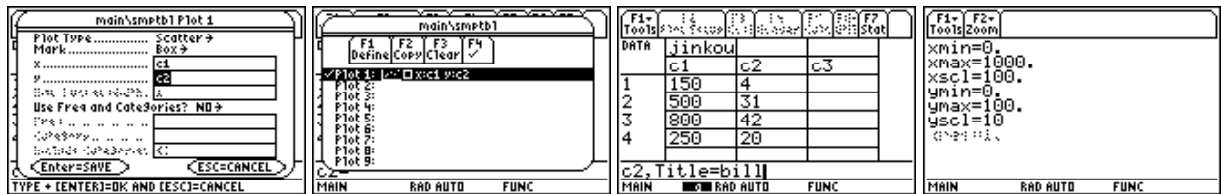
- (5) 表示するグラフのタイプを指定した後は、 x 軸と y 軸にどのデータ列を用いるかを指定します。上段に x 軸のデータ、下段に y 軸のデータを指定します。多数の列にデータを入力しているときは、ここでその列を選択してグラフ化することができます。今取り上げている例では2列しか入力していないので、上段に C1, 下段に C2 を指定します。指定して [ENTER] を押すと、Plot1 の箇所にはチェック印 (✓) がつきます。

なお、グラフのタイプとして「3: Box Plot」や「4: Histogram」を指定した場合は、1つのデータ列だけを指定することになります。

- (6) さらにもう一度 [ENTER] を押して、指定を確定させると、表データに戻ります。
 (7) グラフは [F3] を押すことで表示されますが、その描画範囲に注意する必要があります。今考えているデータの範囲に注意して、 x 軸や y 軸の範囲を適正に指定する必要があります。

x 軸は C1 のデータを指定したので、それは 150~900 の範囲にあります。また、 y 軸は C2 のデータを指定したので、4~73 の範囲にあります。その範囲のデータが表示されるように x, y の範囲を指定します。それは、[F2] で指定することができます。

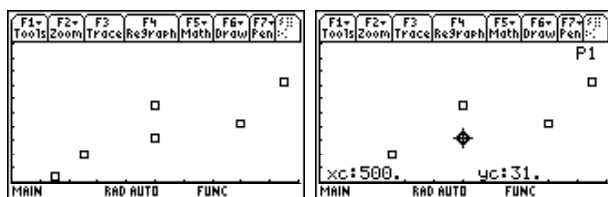
この例では、区切りよく、 $0 \leq x \leq 1000$, $0 \leq y \leq 100$ として、 x の刻み幅 (xscl) は 100, y の刻み幅 (yscl) は 10 としておきます。



列の指定 Plot1 にチェック印 表データに戻る [F2] で範囲指定

- (8) 以上の準備の後で、グラフを表示させるために [F3] を押すと、指定した散布図が表示されます。点が、指定した通り「box」() で表示されていることに注意してください。(4) での指定を変えることで、この記号を、 \times , $+$, \circ などに変更することができます。また、(3) での指定を「2: xyline」に変更すると、これらの点を線で結ぶことができます。

- (9) グラフが表示された状態で [F3] (Trace) を押すと、このグラフ上の点を移動することができます。カーソルが現れて、[Left] [Right] を押すごとに点の間を移動します。表で定義した1行目から順番に移動します。下段には、その値が表示されます。



[F3] でグラフ表示 [F3] でトレース

同様に、点の記号を別な記号に変えたり、直線で結ぶように変更して、グラフ表示のさせ方を理解してください。

9.5 最小2乗法による回帰直線

この例で表示させた散布図をみると、右上がりの傾向が認められることが分かります。実験のデータ処理などでは、このようなときに、データを最もよく近似する直線を求めようとします。それが、最小2乗法による回帰直線です。それは、次のような直線として求められるものです。

今、データを $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ とし、これらを最もよく近似する直線 $y = ax + b$ を求めたいとします。 $x = x_i$ のときの直線上の y 座標 $y = ax_i + b$ と実際の値 y_i との違いを $(y_i - (ax_i + b))^2$ で捉えて、それらの合計を考えます。

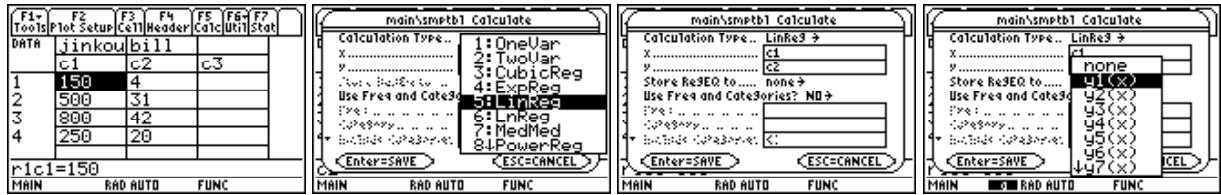
$$\{(\text{実際の値}) - (\text{直線上の値})\}^2 \text{の合計} = \sum_{i=1}^n (y_i - (ax_i + b))^2$$

データが1直線上にあれば、この値は0になります。データが直線から離れていればいるほど、この値は大きな値になります。最小2乗法とは、この値が最小になるような直線 $y = ax + b$ を求める方法です。データ (x_i, y_i) が与えられれば、この差の合計は a, b の関数になります。3年の偏微分法で学ぶ2変数関数の極値を求める方法によれば、この値が最小になる場所があるとすると、それは a, b で偏微分した値が0になる箇所です。未知数が2個で、それぞれで偏微分して0であることから2つの式が得られるので、それを連立方程式として解くことにより a, b の値を求めることができます。数ナビを利用すると、その直線を次のような手順で簡単に求めることができます。

- (1) まず、元の表データに戻りましょう。[APPS]を押してData/Matrix Editorを選択し、直前のデータを利用するので「1: Current」を指定します。
- (2) 表データが表示されたら、[F5]「Calc」を押します。「Calculator Type」で、どのような計算を行うかを指定します。多数の選択肢が表示されますが、次のような内容の計算を行うこととなります。ここでは「5: LinReg」(直線回帰 Line Regression)を選択します。

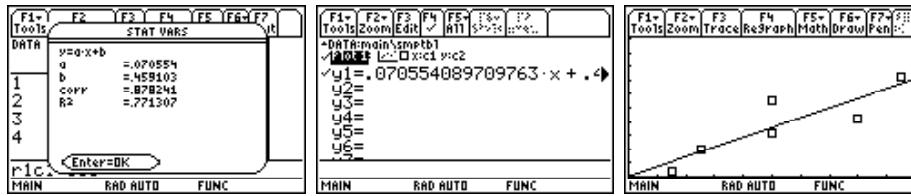
| | |
|-------------|--|
| 1: OneVar | 1つの列のデータに関するいろいろな統計量を計算する。 |
| 2: TwoVar | 2つの列のデータに関するいろいろな統計量を計算する。 |
| 3: CubicReg | データを最もよく近似する3次関数 $y = ax^3 + bx^2 + cx + d$ を求める。 |
| 4: ExpReg | データを最もよく近似する指数関数 $y = ab^x$ を求める。 |
| 5: LinReg | データを最もよく近似する1次関数 $y = ax + b$ を求める。 |
| 6: LnReg | データを最もよく近似する対数関数 $y = a + b \log_e(x)$ を求める。 |
| 7: Logistic | データを最もよく近似する成長曲線 $y = \frac{a}{1 + be^{cx}} + d$ を求める。 |
| 8: MedMed | メジアン-メジアン法をもちいて近似直線 $y = ax + b$ を求める。 |
| 9: PowerReg | データを最もよく近似するべき関数 $y = ax^b$ を求める。 |
| A: QuadReg | データを最もよく近似する2次関数 $y = ax^2 + bx + c$ を求める。 |
| B: QuartReg | データを最もよく近似する4次関数 $y = ax^4 + bx^3 + cx^2 + dx + e$ を求める。 |
| C: SinReg | データを最もよく近似する正弦関数 $y = a \sin(bx + c) + d$ を求める。 |

- 次に、 x 軸と y 軸をどの列にするかを指定します。ここでは、C1 と C2 を指定します。
- 求めた関数 $y = f(x)$ を、関数定義画面に記録するかどうかを、「Store RegEQ to」の箇所指定します。たとえば、 y_1 に記録することになります。



元の表データ 5: LinReg を指定 列を指定 関数を $y_1(x)$ に記録

- 以上を指定して **ENTER** を押すと、結果が表示されます。文字が小さいですが、 $y = ax + b$ の形で式を求めたとき、 $a = 0.070554$ 、 $b = 0.459103$ であることが分かります。「corr」は、5年の確率統計で学ぶ相関係数です。この値が ± 1 に近いほど、相互のデータには直線的な関連性が高いことを示しています。データが1直線上にあるときは、この値は1か-1になります。
- F1** を押すと、関数定義画面の y_1 の箇所に、今求めた式が定義されています。
- F3** を押すと、散布図に直線が描画されます。データをよく近似していることが分かります。

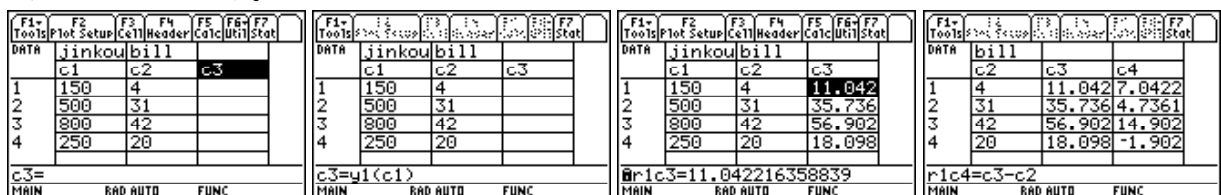


結果の表示 **F1** 関数定義画面 **F3** グラフ表示

9.6 表データでの計算処理

数ナビの表データでは、エクセルと同様、列に対していろいろな計算を行うことができます。前節で求めた回帰直線は、 $y_1(x) = ax + b$ に記録されています。そこで、C3列にその値を出力して実際の値(C2列)と比較してみましょう。

- APPS** を押して Data/Matrix Editor を選択し、「1: Current」を指定して元の表に戻ります。
- 列に対する計算処理は、上から2行目の箇所で指定します。「C3」の箇所で **ENTER** を押してカーソルを入力行に移動させ、 $c_3 = y_1(c_1)$ とすると、C3列に c_1 列の各値に対して $y_1(x) = ax + b$ の値が書き込まれます。
- ENTER** を押すと、C1列に対して $y_1(x) = ax + b$ が計算されてC3列に表示されます。
- 実際の値と $y_1(x)$ で計算された値とがどの程度違うのかをみるため、その差 ($c_3 - c_2$) を C4列に定義します。回帰直線 $y = ax + b$ では、この差の2乗の合計が最小になるように a 、 b の値が定められています。



C3列に式を定義 $c_3 = y_1(c_1)$ C3列の結果 $c_4 = c_3 - c_2$

9.7 いろいろな表操作

挿入と削除

列や行の挿入・削除は、表が表示されている状態で **F6** 「Util」の箇所に登録されています。**2nd** **F1** を押すとメニューが表示されます。

挿入 (insert) を指定すると、挿入するのは個々のセルなのか、行 (row) なのか、列 (column) なのかが問われます。行はカーソルのある行の上側に、列はカーソルのある行の左側に挿入されます。したがって、最初に、挿入したい行や列の場所にカーソルを移動する必要があります。たとえば、C2列の2行目にカーソルを移動して、行や列を挿入したり、それを削除してみてください。いったん削除すると、それを復活する機能はないので注意してください。

並べ換え

列のデータを並べ換えるには、**2nd** **F1** のメニューで「3: Sort Column」を利用します。行の値を保ったまま、カーソルが位置する列の値の順に全体が並べ換えられます。最初に、カーソルを並べ換える列に移動させてから指定すること。そのままでは元の順には復帰できないので注意してください。元に戻すには、新しい列を挿入して、そこに通し番号を入れておく必要があります。通し番号順に並べ換えれば、元に状態に戻すことができます。

列幅の変更

列幅を変更するには、表の出ている状態で **◆** **|** を押します。「Cell Width」の値を変更すると、列幅が変更できます。表の全ての列幅が変更されます。

自動計算

列に対して計算式を定義すると、**ENTER** を押すごとに自動計算されます。複雑な計算式を定義していると、そのたびに時間がかかります。式だけ定義して計算はしないようにするには、「Auto-Calculate」を OFF にします。

The figure consists of four screenshots of a calculator's menu system, each showing a different step in a data manipulation process:

- Row Insertion:** The menu is open to '3: Sort Column'. The user has navigated to '2: row'. The calculator screen shows a table with 4 rows and 3 columns. The values are: Row 1: 150, 4; Row 2: 500, 31; Row 3: 800, 42; Row 4: 250, 20. Below the table, the formula $r2c2=31$ is displayed.
- Insertion Result:** The menu is open to '3: column'. The user has navigated to '2: row'. The calculator screen shows the same table as above, but now with an additional column (c3) inserted to the left of the original c2. The values are: Row 1: 150, c3, 4; Row 2: 500, c3, 31; Row 3: 800, c3, 42; Row 4: 250, c3, 20. Below the table, the formula $r2c2=$ is displayed.
- Column Sort:** The menu is open to '3: Sort Column'. The user has navigated to '3: column'. The calculator screen shows the same table as above, but the values in the c2 column are now sorted in ascending order: Row 1: 150, 4; Row 2: 250, 20; Row 3: 800, 42; Row 4: 500, 31. Below the table, the formula $c1=$ is displayed.
- Sort Result:** The menu is open to '3: Sort Column'. The user has navigated to '3: column'. The calculator screen shows the same table as above, but the values in the c2 column are now sorted in descending order: Row 1: 150, 4; Row 2: 500, 31; Row 3: 800, 42; Row 4: 250, 20. Below the table, the formula $c1=$ is displayed.

F6 (**2nd** **F1**)

行の挿入

挿入結果

並べ換え

このマニュアルは、独立行政法人日本科学技術振興会の平成22年度～24年度科学研究費補助金 基盤研究 (C) 課題番号 22500830 (研究代表者: 梅野善雄) の支援を受けて作成されたものです。