

t 分布理解のための Excel によるシミュレーション¹⁾

— t 分布の確率分布曲線の導出について —

門田 幸太郎ⁱ

Excel によるシミュレーションを使って、 t 分布の確率分布図を導き出す過程を具体的な事象として擬似的に体験することによって、 t 分布および t 検定の原理を直観的に理解することができるようにすることが本稿の目的である。

t 検定は、正規分布をしている母集団から取り出された2つの標本集団の平均値の差が統計的に意味のあるものであるか否かを判定しようというものである。本稿は、 t 検定の原理を、EXCEL によるシミュレーションを利用して直観的に理解できるようにするものである。そのため、1. 標準正規分布している母集団を作成する。次に、2. その中から標本サイズが3からなるすべての小標本を抽出し、その平均値を求める。その後、3. 得られた標本平均値の差の分布を求める。これによって、 t 分布が得られることになる。データから得られた平均値の差が、 t 分布の棄却域に入るか否かを検討することによって、その統計的有意性を判定するのが t 検定であることが示される。

キーワード：Excel, シミュレーション, t 分布, 標準正規分布, NORMDIST 関数, FREQUENCY 関数, INDIRECT 関数

1. 標準正規分布の作成

「 μ を $N(0,1^2)$ に従う確率変数、 $\chi^2(\phi)$ を自由度 ϕ の χ^2 分布に従う確率変数とし、 μ と $\chi^2(\phi)$ が互いに独立であるとしたときに、統計量 $t = \frac{u}{\sqrt{\frac{\chi^2(\phi)}{\phi}}}$ は自由度 ϕ の t 分布に従う。」(守谷栄一, 1987) というのが t 分布の

一般的な定義である。これを理解するには、正規分布はもちろん、 χ^2 分布を理解していないといけない。ここでは、初学者にとっては t 分布がいかなるものなのか理解しがたいところがある。そこで本稿では、EXCEL によるシミュレーションを使って、 t 分布の確率分布図を導き出す過程を具体的な事象として体験することによって、 t 分布および t 検定の原理を直観的に理解することができるようにしようとした。

手続きとしては、まず、母集団となる正規分布を作成する。正規確率密度関数を用いて、各確率変数の値に対応する確率を求める。母集団の全構成要素のサイズを決定するために、得られた確率を定数倍する。一般

i 立命館大学産業社会学部教授, 2015年4月より特別任用教授・立命館大学名誉教授

に、シミュレーションを行う場合、モデルのサイズが大きければ大きいほど好ましいと言えるが、母集団のサイズが大きくなると、標本サイズが小さくても処理量が膨大なものとなる。今もし母集団のサイズが100なら、標本のサイズが3であっても、その組み合わせは ${}_{100}C_3 = 161700$ となる。そこから2組の標本を選び出すとすれば、その組み合わせは ${}_{161700}C_2$ となり、約130兆のデータを処理しなければならない。母集団のサイズを50としても、2兆近いデータを処理しなければならない。これは一般に用いられているPCの処理能力の限界を超えている。こうなると、メモリーが不足したり、処理に長時間が必要となったりして実用性に欠けることになる。大きなデータの活用は今後のCPUの進歩を待たなければならない。ここでは、一般に普及しているPCでの処理の限界と思われる、サイズ28の母集団を想定し、そこからサイズ3の標本を選ぶ組み合わせ ${}_{28}C_3 = 3276$ を作成し、この標本から得られる平均値の中から2組を選び、その組み合わせの ${}_{3276}C_2 = 5364450$ 通りについて検討することにする。本稿のプログラムでは、このような大きなサイズのデータでも対応できるように、データ数を示す変数dを変更することによって対応することができる汎用性を持ったものとなっている。

Excelを利用して、「ND28」と名付けられたシート上に、標準正規分布としてx値が-2.5から2.5にわたる要素の数が28からなる正規分布を作成する。Table 1に示したようにD2セルに-2.5を入力する。次に、D3セルに関数「=D2+0.5」を入力し、その値が2.5になるD12セルまでドラッグしリリースする。これにより、-2.5から0.5のステップで2.5までの数列を得ることができる。

次に、この数列を確率変数として、それぞれの値に対応する確率密度関数を求める。統計関数であるNORMDIST関数を利用する。これは、NORMDIST (X, 平均, 標準偏差, 関数形式) という形式で用いられる。Xは確率密度関数に代入する数値を表し、平均は求める正規分布の平均を、標準偏差は求める正規分布の標準偏差を指定する。関数形式はTRUEとFALSEの論理値がある。TRUEの場合は累積確率を求め、FALSEの場合はXに対応する確率を求めることになる。ここでは標準正規分布を用いるので、平均は0、標準偏差は1とする。D2セルからD12セルに対応する標準正規確率を求めるのでNORMDIST(D2,0,1,FALSE)とする。これは、正規確率密度関数 $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ の独立変数xにD2セルからD12セルの値を、平均を示す μ に0を、標準偏差を示す σ に1を代入することを意味する。

Table 1に示されているように、値が-2.5から2.5にわたる要素の数が28となる正規分布を作成するために、各確率値を14倍する。E2に「=ROUND(NORMDIST(D2,0,1,FALSE)*14,0)」を入力する。

ROUND関数は求める桁数で四捨五入した数値を求めるものである。これはROUND(数値, 桁数) という形式で用いられる。ここでは、NORMDIST(D2,0,1,FALSE)*14を四捨五入して整数値とする。これで各X値に対する個数が求められることになる。それをグラフに表わしたのが、Figure 1である。

Table 1のD列に示された数値がX値に対応し、E列に示された数値がそのX値が現れる回数である頻度を示している。度数分布表がFigure 1に示されている。たとえば、-2.5の出現回数は0回、-2.0は1回、-1.5は2回、-1は3回などである。これらをTable 2に示すように28行にわたって表示する。ここでは0.0を中心として頻度の高いものほど多くなっている。その組み合わせを求める時に、1~28のセル番号のどの3つが選ばれるかは等確率だとして中心値の数ほど多く選ばれるという、平均を中心としたベルシェイプの正規分布の特徴が反映されることになる。これにより、求める正規母集団を得ることができる。標本の要素ごとに毎回独立にサンプリングをした場合、理論的には同一要素が含まれる可能性が考えられるが、ここでは毎回独立にサンプリングをして、標本サイズのデータを入力するのではなく、標本サイズごとに同時にサ

ンプリングを行うことを前提とした。したがって、Table 1 の度数分布表にある階級値 -2.0 のように頻度が 1 の要素が複数回サンプリングされることはなく、階級値 0.0 のように頻度が複数ある要素の場合は複数回サンプリングされることはありうるとした。

Table 1 Frequency distribution table.

X	FREQUENCY	DENSITY
-2.5	0	0.00
-2.0	1	0.04
-1.5	2	0.07
-1.0	3	0.11
-0.5	5	0.18
0.0	6	0.21
0.5	5	0.18
1.0	3	0.11
1.5	2	0.07
2.0	1	0.04
2.5	0	0.00

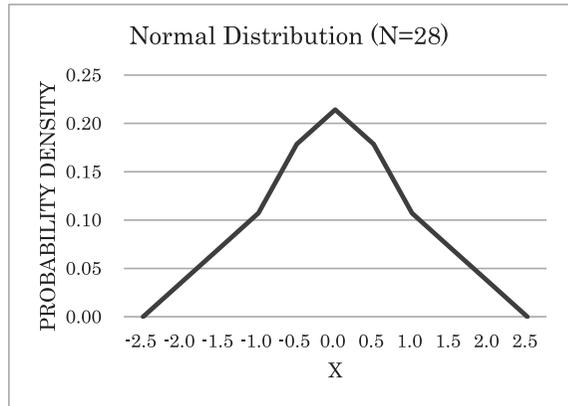


Figure 1 Normal distribution curve.

28

Table 2 Numbers for sampling data.

	A
1	-2.0
2	-1.5
3	-1.5
4	-1.0
5	-1.0
6	-1.0
7	-0.5
8	-0.5
9	-0.5
10	-0.5
11	-0.5
12	0.0
13	0.0
14	0.0
15	0.0
16	0.0
17	0.0
18	0.5
19	0.5
20	0.5
21	0.5
22	0.5
23	1.0
24	1.0
25	1.0
26	1.5
27	1.5
28	2.0
29	

Table 3 The combination of the case of 5 data.

1	2	3
1	2	4
1	2	5
1	3	4
1	3	5
1	4	5
2	3	4
2	3	5
2	4	5
3	4	5

Table 4 The relation between cell numbers and combinations.

(1,1)	(1,2)	(1,3)
1	2	3
(2,1)	(2,2)	(2,3)
1	2	4
(3,1)	(3,2)	(3,3)
1	2	5
(4,1)	(4,2)	(4,3)
1	3	4
(5,1)	(5,2)	(5,3)
1	3	5
(6,1)	(6,2)	(6,3)
1	4	5
(7,1)	(7,2)	(7,3)
2	3	4
(8,1)	(8,2)	(8,3)
2	3	5
(9,1)	(9,2)	(9,3)
2	4	5
(10,1)	(10,2)	(10,3)
3	4	5

2. 標本の抽出

要素の数が28からなる正規分布の中から標本を抽出し、その平均値を求め、その分布を見る。ここでは、例として標本サイズが3からなる小標本を想定し、すべての組み合わせを求め、その平均値を求めることにする。28個のデータから3個の標本を得なければならないが、そのためには ${}_{28}C_3=3276$ （個）の組み合わせを求めなければならない。3個のデータの組み合わせを求めるプログラムとして、Excelに備え付けられたVBA（Visual Basic for Applications）コードを利用する。ここでは説明の便宜上、データ数を5として組み合わせを求める全過程を取り上げることにする。求めるべきすべての組み合わせはTable 3のようになる。この組み合わせを3桁の数字と考えると、123から始まり、右端の1の位が5まで増えると中央の10の位の2が3となる。ここでも1の位が5まで増えると10の位の3が4となる。10の位が4となると1の位は5となり、次は左端の100の位が2となり、234となる。1の位は5となると235から245となり、最後は1の位と10の位が同時に上がって345となり、これが最後となる。このように5つの数の中から3つの数の組み合わせを選び、3桁の昇順の整数を作成する方法を、VBAを用いてプログラミングする。

VBAを利用することができるようにするには、Figure 2に示すようにExcelのメニューにある「開発」タブをクリックし、「コード」オプションの中の「Visual Basic」ボタンを選択することになる。ツールバーにあるのアイコンにより「ユーザーフォームの挿入」を選択するとツールボックスが現れて、フォームの編集が可能になる。Figure 3に示すように、ツールボックスからコマンドボタンを取り込み、プロパティのCaptionを「28C3」と変えてから、ダブルクリックするとコード・ウインドが現れてコード編集が可能になる。これにより、「28C3」のコマンドボタンをクリックすることによりFigure 4に示すように「Private Sub CommandButton1_Click0」と「End Sub」のコードが実行されるプログラムを作ることができる。実現しようとする結果とその数字の収納場所を示すセル番号との関係を示すとTable 4のようになる。

Figure 5のフローチャートとFigure 6のプログラムコード、Table 5の処理過程表にしたがってプログラムを説明する。k, l, mの組み合わせを3桁の整数とみなした場合、100位となる数字をk, 10位となる数字をl, 1位となる数字をmとし、初期値としてkに1, lに2, mに3を代入する。セル位置を示す変数としてi, jを用い、初期値としてはともに1を入れる。データ数を示す変数としてdを用いる。ここではデータ数を5とする。ラベルTop:は繰り返し作業の制御のために使われる。Cells(i, j)=kは(1,1)のセルに1を入れる。kは1の状態なので、 $k > d - 2$ の条件式は $1 > 5 - 2$ となり成立しないので通過する。次のセル(1,2)に移動するため $j = j + 1$ とする。右辺のjに1が入り、 $1 + 1$ で左辺のjには2が入る。Cells(i, j)=1により、(1,2)のセルに1の初期値2が入る。2番目の $j = j + 1$ では、右辺のjには2が入り、左辺のjは3となる。Cells(i, j)=mには(1,3)のセルにmの初期値3が入る。m=m+1では、右辺のmには3が入り、左辺のmは4となる。mの4はdの5より小さいので、1の2とd-1の4もIf文の条件式 $1 > d - 1$ に合わないのでEnd Ifに飛ぶ。次の条件式 $k > d - 2$ では1と5-2を比較することになるので条件式は成立しない。i=i+1は次の行のセルに移るための式であり、右辺のiには1が入り、左辺のiは2となる。j=j-2は右辺のjには3が入り、左辺のjは1となる。これにより、(1,3)のセルから(2,1)のセル位置に移動することになる。GoTo TopでラベルTop:に移動することになる。

再び、Top:に戻り、Cells(i, j)=kは(2,1)のセルに1を入れる。条件式 $k > d - 2$ は $1 > 3$ となり、成立しないのでパスとなる。j=j+1は $2 = 1 + 1$ となり、左辺のjは2となる。Cells(i, j)=1は(2,2)のセルに1の初期値の

ままの 2 が入る。j=j+1 で左辺の j には 3 が入るので、Cells(i,j) = m は (2,3) セルに m=m+1 により得られた 5 が入る。

m>d は 5>5 となり、条件式は成立しないので End If までパスすることになる。次の If 文の条件式 1>d-1

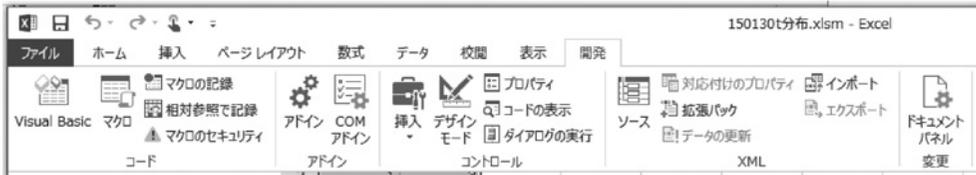


Figure 2 Usage of VBA.

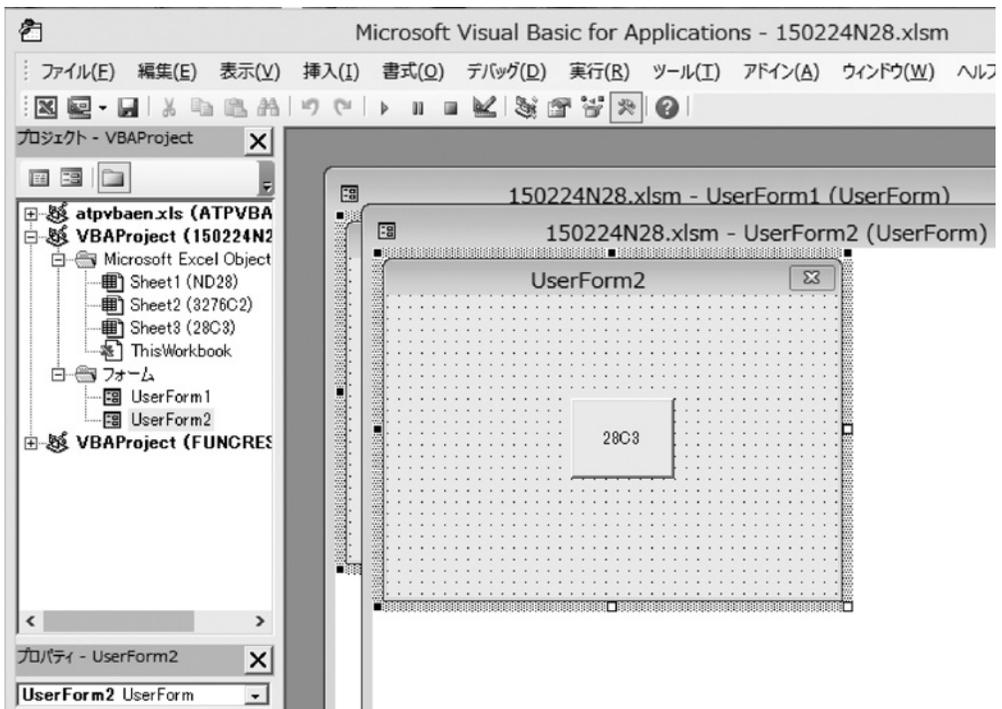


Figure 3 Form of VBA.

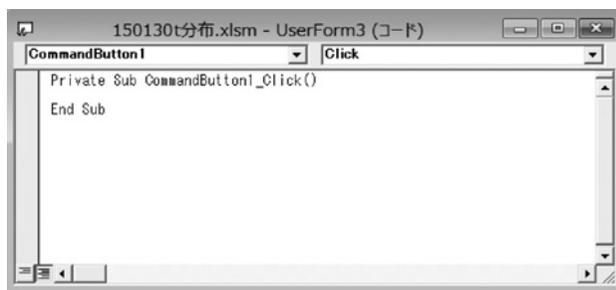


Figure 4 Code window of VBA.

は $2 > 5 - 1$ となるので成立しないので、ここでも End If までパスすることになる。続く If 文の条件式 $k > d - 2$ は $1 > 5 - 2$ となるので成立しないのでこの If 文もパスすることになる。次の 2 行 $i = i + 1$ と $j = j - 2$ は i に 3, j に 1 が入る。

3 回目に Top: に戻り, $\text{Cells}(i, j) = k$ は $\text{Cells}(3, 1) = 1$ となる。 $k > d - 2$ は $1 > 5 - 2$ は成立しないのでパスされる。 $j = j + 1$ で左辺の j は 2 となり, $\text{Cells}(i, j) = 1$ は $\text{Cells}(3, 2) = 2$ となる。 j が 1 増えて, $\text{Cells}(i, j)$ は $\text{Cells}(3, 3)$ となり, $\text{Cells}(i, j) = m$ は $\text{Cells}(3, 3) = 5$ となる。 $m = m + 1$ は右辺の m が 5 なので, 左辺の m には 6 が入る。次の If 文は条件式 $m > d$ は $6 > 5$ となり成立するので, $l = l + 1$ と $m = m + 1$ が実行される。 $l = l + 1$ は右辺の l に 2 が入り, 左辺の l は 3 となる。 $m = m + 1$ は右辺の l が 3 なので, 左辺の m には 4 が入る。条件式 $l > d - 1$ は $3 > 5 - 1$ で成立しないので, End If まで飛ぶ。条件式 $k > d - 2$ は $1 > 5 - 2$ は成立しないので, パスされる。 i は 4 となり, j が 1 に戻る。

4 回目に Top: に戻り, $\text{Cells}(i, j) = k$ は $\text{Cells}(4, 1) = 1$ となる。 $k > d - 2$ は $1 > 5 - 2$ は成立しないのでパスされる。 j は 2 となり, $\text{Cells}(i, j) = 1$ は $\text{Cells}(4, 2) = 3$ となる。 j が 1 増えて, $\text{Cells}(i, j)$ は $\text{Cells}(4, 3)$ となり, $\text{Cells}(i, j) = m$ は $\text{Cells}(4, 3) = 4$ となる。 $m = m + 1$ は右辺の m が 4 なので, 左辺の m には 5 が入る。次の If 文は条件式 $m > d$ は $5 > 5$ となり成立しない。条件式 $l > d - 1$ は $3 > 5 - 1$ で成立しないので, End If まで飛ぶ。条件式 $k > d - 2$ は $1 > 5 - 2$ は成立しないので, パスされる。 i は 5 となり, j は 1 に戻る。

5 回目に Top: に戻り, $\text{Cells}(i, j) = k$ は $\text{Cells}(5, 1) = 1$ となる。 $k > d - 2$ は $1 > 5 - 2$ は成立しないのでパスされる。 j は 2 となり, $\text{Cells}(i, j) = 1$ は $\text{Cells}(5, 2) = 3$ となる。 j が 1 増えて, $\text{Cells}(i, j)$ は $\text{Cells}(5, 3)$ となり, $\text{Cells}(i, j) = m$ は $\text{Cells}(5, 3) = 5$ となる。 $m = m + 1$ は右辺の m が 5 なので, 左辺の m には 6 が入る。次の If 文は条件式 $m > d$ は $6 > 5$ で成立するので, $l = l + 1$ と $m = m + 1$ が実行される。 $l = l + 1$ は右辺の l に 3 が入り, 左辺の l は 4 となる。 $m = m + 1$ は右辺の l が 4 なので, 左辺の m には 5 が入る。条件式 $l > d - 1$ は $4 > 5 - 1$ で成立しないので, End If まで飛ぶ。条件式 $k > d - 2$ は $1 > 5 - 2$ は成立しないので, パスされる。 i は 6 となり, j は 1 に戻る。

6 回目に Top: に戻り, $\text{Cells}(i, j) = k$ は $\text{Cells}(6, 1) = 1$ となる。 $k > d - 2$ は $1 > 5 - 2$ は成立しないのでパスされる。 j は 2 となり, $\text{Cells}(i, j) = 1$ は $\text{Cells}(6, 2) = 4$ となる。 $\text{Cells}(i, j) = m$ は $\text{Cells}(6, 3) = 5$ となる。 $m = m + 1$ は右辺の m が 5 なので, 左辺の m には 6 が入る。次の If 文の条件式 $m > d$ は $6 > 5$ となりで成立するので, $l = l + 1$ と $m = m + 1$ が実行される。 $l = l + 1$ は右辺の l に 4 が入り, 左辺の l は 5 となる。 $m = m + 1$ は右辺の l に 5 が入り, 左辺の m は 6 となる。条件式 $l > d - 1$ は $5 > 5 - 1$ で成立するので, $k = k + 1$ と $l = k + 1$ と $m = m + 1$ が実行される。 $k = k + 1$ は右辺の k が 1 なので, 左辺の k は 2 となる。 $l = k + 1$ は右辺の k が 2 なので, 左辺の l は 3 となる。 $m = m + 1$ は右辺の l が 3 なので, 左辺の m は 4 となる。条件式 $k > d - 2$ は $2 > 5 - 2$ は成立しないので, パスされる。 i は 7 となり, j は 1 に戻る。

7 回目に Top: に戻り, $\text{Cells}(i, j) = k$ は $\text{Cells}(7, 1) = 2$ となる。 $k > d - 2$ は $2 > 5 - 2$ は成立しないのでパスされる。 j は 2 となり, $\text{Cells}(i, j) = 1$ は $\text{Cells}(7, 2) = 3$ となる。 $\text{Cells}(i, j) = m$ は $\text{Cells}(7, 3) = 4$ となる。 $m = m + 1$ は右辺の m が 4 なので, 左辺の m には 5 が入る。次の If 文の条件式 $m > d$ は $5 > 5$ となりで成立しない。条件式 $l > d - 1$ は $3 > 5 - 1$ で成立しないので, End If まで飛ぶ。条件式 $k > d - 2$ は $2 > 5 - 2$ は成立しないので, パスされる。 i は 8 となり, j は 1 に戻る。

8 回目に Top: に戻り, $\text{Cells}(i, j) = k$ は $\text{Cells}(8, 1) = 2$ となる。 $k > d - 2$ は $2 > 5 - 2$ は成立しないのでパスされる。 j は 2 となり, $\text{Cells}(i, j) = 1$ は $\text{Cells}(8, 2) = 3$ となる。 $\text{Cells}(i, j) = m$ は $\text{Cells}(8, 3) = 5$ となる。 $m = m + 1$ は右辺の m が 5 なので, 左辺の m には 6 が入る。次の If 文の条件式 $m > d$ は $6 > 5$ となりで成立する。 $l = l + 1$

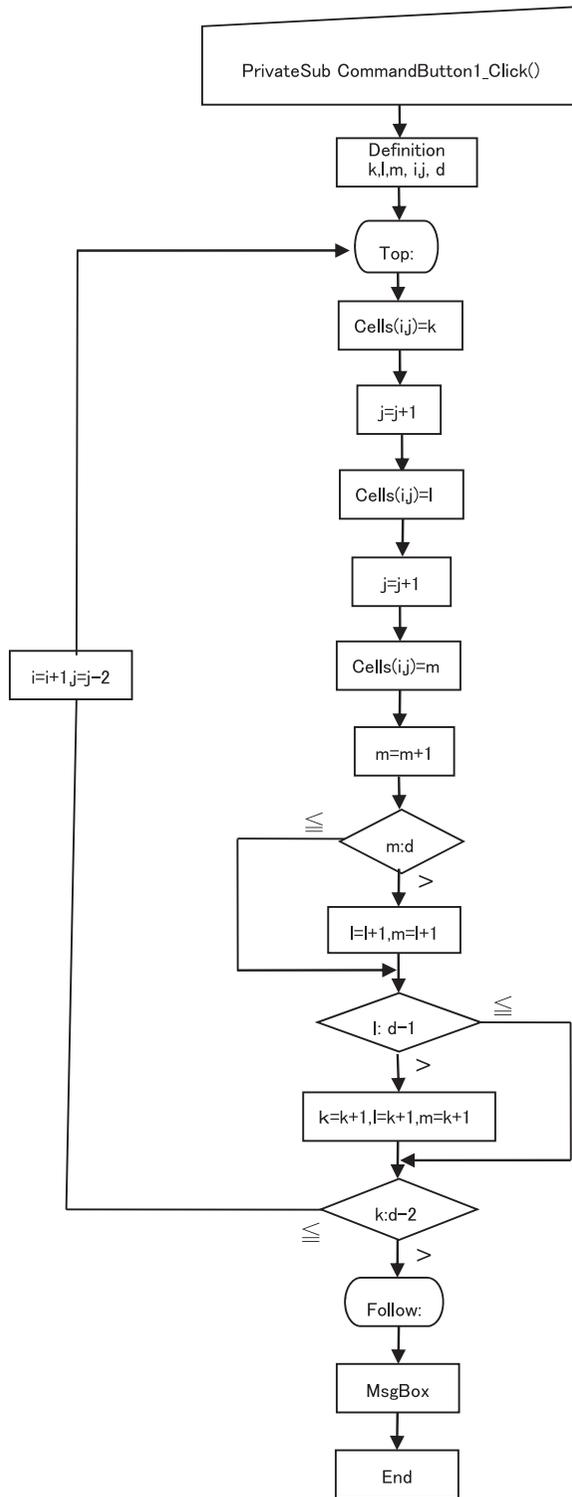


Figure 5 Flowchart for the combination of three numbers.

と $m=1+1$ が実行される。 $l=1+1$ は右辺の l に 3 が入り、左辺の l は 4 となる。 $m=1+1$ は右辺の l に 4 が入り、左辺の m は 5 となる。条件式 $l>d-1$ は $4>5-1$ で成立しないので、End If まで飛ぶ。条件式 $k>d-2$ は $2>5-2$ は成立しないので、パスされる。 i は 9 となり、 j は 1 に戻る。

9 回目に Top: に戻り、Cells(i,j)= k は Cells($9,1$)= 2 となる。 $k>d-2$ は $2>5-2$ は成立しないのでパスされる。 j は 2 となり、Cells(i,j)= 1 は Cells($9,2$)= 4 となる。Cells(i,j)= m は Cells($9,3$)= 5 となる。 $m=m+1$ は右辺の m が 5 なので、左辺の m には 6 が入る。次の If 文の条件式 $m>d$ は $6>5$ となりで成立する。 $l=1+1$ と $m=1+1$ が実行される。 $l=1+1$ は右辺の l に 4 が入り、左辺の l は 5 となる。 $m=1+1$ は右辺の l に 5 が入り、左辺の m は 6 となる。条件式 $l>d-1$ は $5>5-1$ で成立する。 $k=k+1$ と $l=k+1$ と $m=1+1$ が実行される。 $k=k+1$ は右辺の k が 2 なので、左辺の k は 3 となる。 $l=k+1$ は右辺の k が 3 なので、左辺の l は 4 となる。 $m=1+1$ は右辺の l が 4 なので、左辺の m は 5 となる。条件式 $k>d-2$ は $3>5-2$ となり成立しないので、パスされる。 i は 10 となり、 j は 1 に戻る。

10 回目に Top: に戻り、Cells(i,j)= k は Cells($10,1$)= 3 となる。 $k>d-2$ は $3>5-2$ は成立しないのでパスされる。 j は 2 となり、Cells(i,j)= 1 は Cells($10,2$)= 4 となる。Cells(i,j)= m は Cells($10,3$)= 5 となる。 $m=m+1$ は右辺の m が 5 なので、左辺の m には 6 が入る。次の If 文の条件式 $m>d$ は $6>5$ となりで成立する。 $l=1+1$ と $m=1+1$ が実行される。 $l=1+1$ は右辺の l に 4 が入り、左辺の l は 5 となる。 $m=1+1$ は右辺の l に 5 が入り、左辺の m は 6 となる。

条件式 $l>d-1$ は $5>5-1$ で成立する。 $k=k+1$ と $l=k+1$ と $m=1+1$ が実行される。 $k=k+1$ は右辺の k が 3 なので、左辺の k は 4 となる。 $l=k+1$ は右辺の k が 4 なので、左辺の l は 5 となる。 $m=1+1$ は右辺の l が 5 なので、左辺の m は 6 となる。条件式 $k>d-2$ は $4>5-2$ となり成立するので、ラベル Follow: に移動する。MsgBox (“終了します。”) はプログラムが突然終了しないように、「終了します。」というメッセージを持つメッセージ・ボックスを表示する。

コードの編集が完了しプログラムの保存が完了したならば、「開発」をクリックし「Visual Basic」でツールバーにある実行ボタン  をクリックすれば 1 から 28 までの中から選ばれた 3 つの数字の組み合わせが表示される。ただし、実行に当たっては、 d の値を 5 から 28 に変更しておく必要がある。

実行の結果、「28C3」と名付けられたシートの 1 行目の「1 2 3」から 3276 行目の「26 27 28」まで 3 数の組み合わせが表示される。Table 6 に実行結果の一部を表示する。「28C3」シート上の A 列、B 列、C 列の数字に対応する「ND28」シートの数値が Table 2 の A 列の値から選ばれることになる。たとえば、Table 6 の 1 行目の 1 、 2 、 3 に対応する Table 2 (「ND28」シート) の -2.0 (A1)、 -1.5 (A2)、 -1.5 (A3) にそれぞれ対応しており、Table 6 の 3276 行目の 26 、 27 、 28 は 1.5 (A26)、 1.5 (A27)、 2.0 (A28) に対応している。そのため「28C3」シート上から「ND28」シートのセル内容を参照できる形式に変更してやらなければならない。「28C3」シートの E 1 セルに関数「="ND28!A" & A1」を入力する。「ND28!A" & A1 は「ND28」シートの A 列のセルから A1 に示された数値の行の内容を選び出すことを意味している。入力後、E1 の内容を G1 までコピーする。次に I1 に関数「=INDIRECT(E1)」を入力する。INDIRECT() は () 内に示されたセルの内容を間接参照することを意味する。INDIRECT(E1) は E1 に示された「ND28!A1」の内容である -2.0 を示すことになる。I1 の内容を J1 と K1 にコピーする。その後、E1 から K1 を「26 27 29」の組み合わせが表示されている 3276 行までコピーする。ここまでの手続きで、Table 1 で示された -2.5 から 2.5 にわたる要素の数が 28 からなる正規分布からデータ数が 3 の組み合わせが得られたことになる。

Table 6 に示したように、I 列、J 列、K 列に示された 3 つの標本の平均を求める。L1 セルに平均を求めるた

```
Private Sub CommandButton1_Click()
```

```
    Dim k, l, m As Long
```

```
    Dim i, j As Long
```

```
    Dim d As Long
```

```
    k = 1: l = 2: m = 3
```

```
    i = 1: j = 1
```

```
    d = 5 'd データ数
```

```
Top:
```

```
    Cells(i, j) = k
```

```
    j = j + 1
```

```
    Cells(i, j) = l
```

```
    j = j + 1
```

```
    Cells(i, j) = m
```

```
    m = m + 1
```

```
    If m > d Then
```

```
        l = l + 1
```

```
        m = l + 1
```

```
    End If
```

```
    If l > d - 1 Then
```

```
        k = k + 1
```

```
        l = k + 1
```

```
        m = l + 1
```

```
    End If
```

```
    If k > d - 2 Then GoTo Follow
```

```
    i = i + 1
```

```
    j = j - 2
```

```
    GoTo Top
```

```
Follow:
```

```
    MsgBox ("終了します。")
```

```
End
```

```
End Sub
```

Figure 6 Code for the combination of three numbers.

めの AVERAGE 関数を用いて「= AVERAGE(I1,K1)」と入力した後, L3276 セルまでコピーする。これにより L 列に 3 つの標本の平均を求めることができる。この平均が示す分布も自由度が 2 の t 分布の一種となっている。

この分布の様子を見るため, 度数分布を求めるための関数である FREQUENCY 関数を用いる。度数分布の階級を作成するため, N2 セルに「-2.5」と入力した後, N3 セルには「=N2+0.5」と入力し, その後 N12 までドラッグしてリリースする。これで -2.5 から +2.5 までの度数の階級が作成される。次に, L1 から L3276 までの度数分布を得るために O2 セルに「=FREQUENCY(L1:L3276,N2:N12)」を入力する。FREQUENCY 関数は「FREQUENCY 関数 (対象範囲, 階級の区切りとなる数値)」という形式で用いられる。O2 セルに入力した後, O2 セルから O12 セルをフォーカスした状態でツールバーの関数入力欄をクリックして, 範囲指定の文字が, 対象範囲が青色で階級値が赤色になっている時に Ctrl+Shift+Enter と 3 つのキーを同時に押す。すると O2 セルから O12 セルの関数が「{=FREQUENCY(L1:L3276,N2:N12)}」となり, 各階級に対応した度数が表示される (Table 7 参照)。O14 には全度数を求めるため「=SUM(O2:O12)」が入れている。この分布の様子を図示するため, O2 セルから O12 セルまでをフォーカスした状態で「挿入」タブの「グラフ」オプションの中から「折れ線グラフの挿入」ボタンをクリックする。

横軸の表示を配列 O2:O12 に変更するため, グラフの横軸ラベル領域をクリックする。とすると, Figure 7 の

Table 7 Example of means of 3 numbers and frequencies table of means.

X	FREQUENCY	DENSITY
-2.5	0	0.00
-2.0	0	0.00
-1.5	7	0.00
-1.0	129	0.04
-0.5	589	0.18
0.0	1113	0.34
0.5	981	0.30
1.0	396	0.12
1.5	60	0.02
2.0	1	0.00
2.5	0	0.00

3276



Figure 7 The choice of data source.

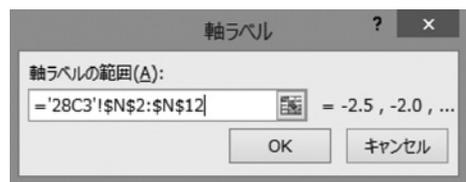


Figure 8 Change of Axis label.

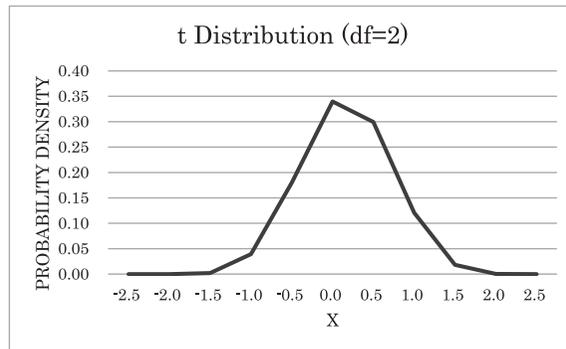


Figure 9 Example of t-distribution.

ような「データソースの選択」が表示される。その中の「軸（項目）軸ラベル」の「編集」をクリックすると Figure 8 のように軸ラベルの範囲指定ができるようになる。ここで N2:N12 のリストを選べば横軸の表示が変更できる。その結果得られた分布図が Figure 9 に示された t 分布である。この場合、3 つの要素からなる標本の平均を取り上げているので、自由度は 2 となる。

3. 標本平均の差の分布

今までの過程で、「1. 標準正規分布の作成」の最後で述べた抽出方法によって、正規分布する 28 の数値から抽出された 3 つの値の平均値が ${}_{28}C_3 = 3276$ 個得られた。これらの平均の間で可能な組み合わせの分布も t 分布していることを示す。その組み合わせは ${}_{3276}C_2 = 5364450$ 個ある。Excel のワークシートは最大 1048576 行 × 16384 列なので、最大行を利用しても、少なくとも 6 列に及ぶことになる。それゆえ、一定の表示行数ごとに列を変えて表示する工夫をしなければならない。ここでは、プログラムコードを説明のための例として、5 個のデータから 3 つを選ぶことによって得られた ${}_5C_3 = 10$ の組み合わせの中から 2 つを選ぶ場合を取り上げる。10 個の中から 2 つを選び組み合わせは ${}_{10}C_2 = 45$ 通りある。これを 10 行ごとに表示する方法を説明する。コードのロジックはデータ数と表示行数を換えてもまったく同じである。

Figure 10 のフローチャートと Figure 11 のコードと Table 10 の処理過程表をもとに説明する。実行結果を Table 8 に、セル番号とセルの内容との関係を Table 9 に示す。

プログラムコードは、まず変数の定義で始まる。k, l と i, j と d, v が長整数型 long として定義されている。通常、整数の変数は Integer として定義される。これはメモリーサイズが 2 byte で $(2^8)^2 = 65536$ となり、 $-32768 \sim 32767$ の範囲の整数しか扱えない。これから扱おうとするデータ数はこの範囲を超えているので、メモリーサイズが 4 byte で $(2^8)^4 = 42.9$ 兆となり、 ± 21 兆を超える範囲の整数を扱うことのできるようにするため上記の諸変数を長整数とする。k と l は変数の値を示し、i と j はそれぞれの k と l が割り当てられるセルのアドレスを示している。d は扱うデータ数を示し、ここでは 10 とする。v は表示する行数を示し、ここでは 10 としている。

変数定義ののち、Top: を通過し、初期値にしたがって Cells(1,1) つまり A1 セルに k の初期値 1 が入る。表示が 9 行を超えているかどうかを判定し、列位置 j を 1 だけ増やして次の列に表示できるようにする。Cells(1,2) つまり B1 セルに 1 の初期値 2 が入る。1 を 1 増やした後、1 とデータ数 d を比較し、3 と 10 を比較す

Table 8 Example of means of 2 numbers.

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
2	1	2					2	4					3	7					5	6					7	9		
3	1	3					2	5					3	8					5	7					7	10		
4	1	4					2	6					3	9					5	8					8	9		
5	1	5					2	7					3	10					5	9					8	10		
6	1	6					2	8					4	5					5	10					9	10		
7	1	7					2	9					4	6					6	7								
8	1	8					2	10					4	7					6	8								
9	1	9					3	4					4	8					6	9								
10	1	10					3	5					4	9					6	10								
11	2	3					3	6					4	10					7	8								

Table 9 The relation between cell numbers and combination ${}_{10}C_2$.

(1,1)	(1,2)			(1,7)	(1,8)			(1,13)	(1,14)			(1,19)	(1,20)			(1,25)	(1,2)
1	2			2	4			3	7			5	6			7	9
(2,1)	(2,2)			(2,7)	(2,8)			(2,13)	(2,14)			(2,19)	(2,20)			(2,25)	(2,2)
1	3			2	5			3	8			5	7			7	10
(3,1)	(3,2)			(3,7)	(3,8)			(3,13)	(3,14)			(3,19)	(3,20)			(3,25)	(3,2)
1	4			2	6			3	9			5	8			8	9
(4,1)	(4,2)			(4,7)	(4,8)			(4,13)	(4,14)			(4,19)	(4,20)			(4,25)	(4,2)
1	5			2	7			3	10			5	9			8	10
(5,1)	(5,2)			(5,7)	(5,8)			(5,13)	(5,14)			(5,19)	(5,20)			(5,25)	(5,2)
1	6			2	8			4	5			5	10			9	10
(6,1)	(6,2)			(6,7)	(6,8)			(6,13)	(6,14)			(6,19)	(6,20)				
1	7			2	9			4	6			6	7				
(7,1)	(7,2)			(7,7)	(7,8)			(7,13)	(7,14)			(7,19)	(7,20)				
1	8			2	10			4	7			6	8				
(8,1)	(8,2)			(8,7)	(8,8)			(8,13)	(8,14)			(8,19)	(8,20)				
1	9			3	4			4	8			6	9				
(9,1)	(9,2)			(9,7)	(9,8)			(9,13)	(9,14)			(9,19)	(9,20)				
1	10			3	5			4	9			6	10				
(10,1)	(10,2)			(10,7)	(10,8)			(10,13)	(10,14)			(10,19)	(10,20)				
2	3			3	6			4	10			7	8				

ることになるので、不等号 \leq が成立するので、 $i=i+1$ と $l=j-1$ により、 i には2が、 l には1が入り、次のセルのアドレス(2,1)が用意される。 i と v は2と10を比較することになるので2回目のTop:へもどる。Cells(2,1)すなわちA2セルが1となり、 j が2となる。Cells(2,2)すなわちB2セルが3となる。1を4にして、 l と d を比較する。 $4 \leq 10$ なので、 i を3、 j を1として3回目のTop:へもどる。Cells(3,1)に1が、Cells(3,2)に4が入る。この過程を8回繰り返えし、9回目にTop:を通過するとき、 $k=1$ 、 $l=10$ 、 $i=9$ 、 $j=1$ の状態を通過する。Cells(9,1)のA9セルに1を入れ、Cells(9,2)のB9セルに10を入れる。この時、 $l:d$ の比較は11:10となり、If $l>d$ Then 文の条件式が成り立つことになる。そのため、 $k=k+1$ と $l=k+1$ が実行される。これにより k は2、 l は3となる。これは求める組み合わせが(1,10)から(2,3)に変化することを意味している。この時、 j の値に6が追加されるのは表示する列を変えるためのものである。20回、30回、40回目のTop:の通過の時も同様である。すべての処理過程はTable 10に示したとおりである。

3276個の平均値の差を取り扱う場合、そこから2つを選ぶことになる。そのため、前述のように5364450の組み合わせを扱わなければならない。これをプログラムで実行する場合、データ数 $d=3276$ 、表示行数 $v=1000000$ とする。実行結果の一部をTable 11に示す。2つの数字がA、B列とG、H列、M、N列、S、T列、Y、Z列、AE、AF列、の6組に表示されている。その数字に対応する「28C3」シートのL列にある数値を取り込む、たとえば、C1セルにはA1セルに収納されている数値1に対応するTable 6のL1セルの値-1.6667を取り込まなければならない。そのため、Table 11の「3276C2」シート上のC1セルに関数「=INDIRECT("28C3!L" & A1)」を入力する。これは、前述したINDIRECT関数で、ここでは、「28C3」シートのL列にある「3276C2」シート上のA1セルにある内容、つまり、「28C3」シートのL1セル (Table 6)に

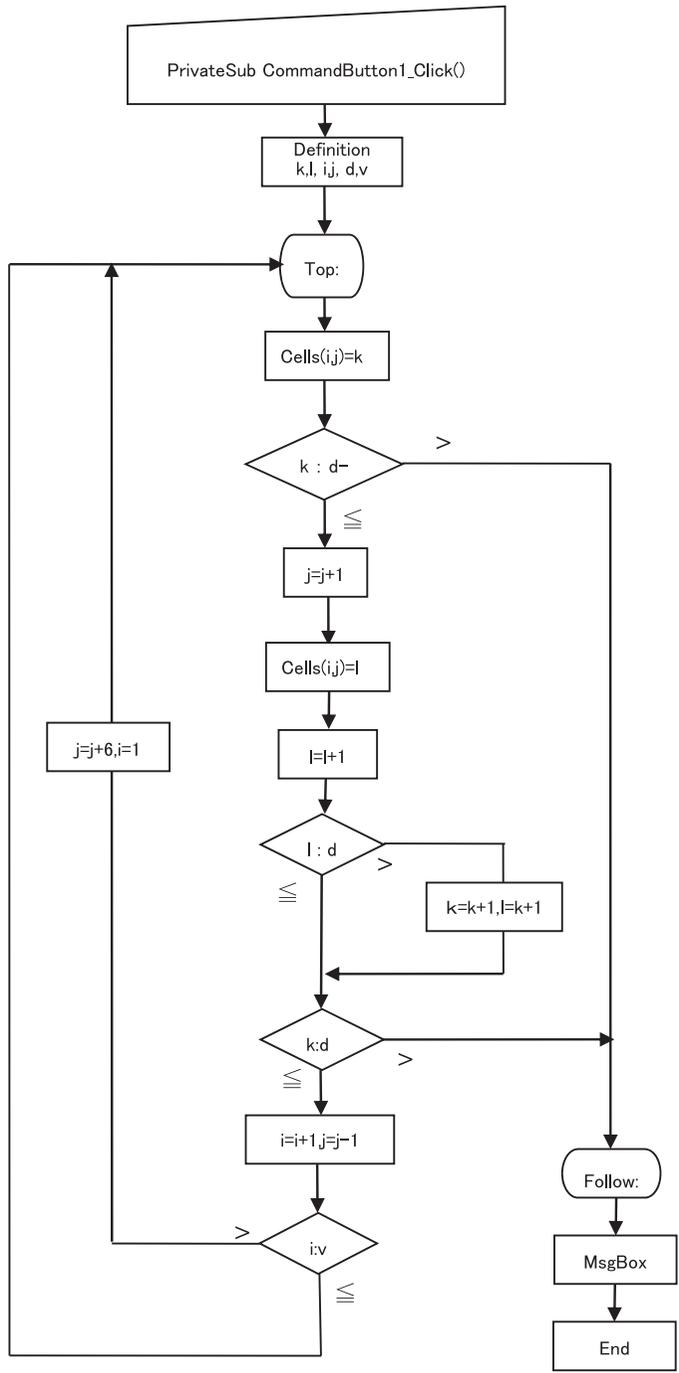


Figure 10 Code for the combination of three numbers.

```
Private Sub CommandButton1_Click()
```

```
    Dim k, l As Long
```

```
    Dim i, j As Long
```

```
    Dim d, v As Long
```

```
    k = 1: l = 2
```

```
    i = 1: j = 1
```

```
    d = 10 'd データ数
```

```
    v = 10 '表示行数
```

```
Top:
```

```
    Cells(i, j) = k
```

```
    If k > d - 1 Then GoTo Follow
```

```
    j = j + 1
```

```
    Cells(i, j) = l
```

```
    l = l + 1
```

```
    If l > d Then
```

```
        k = k + 1
```

```
        l = k + 1
```

```
    End If
```

```
    If k > d - 1 Then GoTo Follow
```

```
    i = i + 1
```

```
    j = j - 1
```

```
    If i > v Then
```

```
        j = j + 6
```

```
        i = 1
```

```
    End If
```

```
    GoTo Top
```

```
Follow:
```

```
    MsgBox ("終了します。")
```

```
End
```

```
End Sub
```

Figure 11 Code for the combination of three numbers.

ある-1.66667を参照することを意味する。D1セルには関数「=INDIRECT("28C3IL" & B1)」が入力されており、Table 6のL2にある-1.5を参照している。E1セルには「=C1-D1」が入力されている。これはC列とD列との対応する行の差を求めるものである。このC1からE1をフォーカスした状態で1000000行目までコピーする。以下、I~K列、O~Q列、U~W列、AA~AC列、AG~AF列も同様にコピーしていく、ただしAG~AF列だけは364450行目までである。これにより、E列、K列、Q列、W列、AC列、AI列にサイズ3の標本の平均の差が5364450個得られることになる。

次に、FREQUENCY関数を用いて、これらの標本平均の差の分布を検討する。ここでの、FREQUENCY関数は度数を求める対象範囲が複数にわたっているので、前述の対象範囲の指定の仕方とは異なるものになる。

Table 10 The processing of the program for 10C2. (1/3)

k	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2
l	2	3	4	5	6	7	8	9	10	3	4	5	6	7	8
i	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5
j	1	1	1	1	1	1	1	1	1	1	7	7	7	7	7
Top:	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	13th	14th	15th
Cells(i,j)=k	(1,1)=1	(2,1)=1	(3,1)=1	(4,1)=1	(5,1)=1	(6,1)=1	(7,1)=1	(8,1)=1	(9,1)=1	(10,1)=2	(1,7)=2	(2,7)=2	(3,7)=2	(4,7)=2	(5,7)=2
k : d-1	1:10-1	1:10-1	1:10-1	1:10-1	1:10-1	1:10-1	1:10-1	1:10-1	1:10-1	2:10-1	2:10-1	2:10-1	2:10-1	2:10-1	2:10-1
≤	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
j = j#1	2=1+1	2=1+1	2=1+1	2=1+1	2=1+1	2=1+1	2=1+1	2=1+1	2=1+1	2=1+1	8=7+1	8=7+1	8=7+1	8=7+1	8=7+1
Cells(i,j)=l	(1,2)=2	(2,2)=3	(3,2)=4	(4,2)=5	(5,2)=6	(6,2)=7	(7,2)=8	(8,2)=9	(9,2)=10	(10,2)=3	(1,8)=4	(2,8)=5	(3,8)=6	(4,8)=7	(4,8)=8
l = l#1	3=2+1	4=3+1	5=4+1	6=5+1	7=6+1	8=7+1	9=8+1	10=9+1	11=10+1	4=3+1	5=4+1	6=5+1	7=6+1	8=7+1	9=8+1
l : d	3:10	4:10	5:10	6:10	7:10	8:10	9:10	10:10	11:10	4:10	5:10	6:10	7:10	8:10	9:10
>	×	×	×	×	×	×	×	×	○	×	×	×	×	×	×
* k = k+1	↓	↓	↓	↓	↓	↓	↓	↓	↓	2=1+1	↓	↓	↓	↓	↓
* l = k+1	↓	↓	↓	↓	↓	↓	↓	↓	↓	3=2+1	↓	↓	↓	↓	↓
k : d-1	1:10-1	1:10-1	1:10-1	1:10-1	1:10-1	1:10-1	1:10-1	1:10-1	2:10-1	2:10-1	2:10-1	2:10-1	2:10-1	2:10-1	2:10-1
≤	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
i = i#1	2=1+1	3=2+1	4=3+1	5=4+1	6=5+1	7=6+1	8=7+1	9=8+1	10=9+1	11=10+1	2=1+1	3=2+1	4=3+1	5=4+1	5=4+1
j = j-1	1=2-1	1=2-1	1=2-1	1=2-1	1=2-1	1=2-1	1=2-1	1=2-1	1=2-1	1=2-1	7=8-1	7=8-1	7=8-1	7=8-1	7=8-1
i : v	2:10	3:10	4:10	5:10	6:10	7:10	8:10	9:10	10:10	11:10	2:10	3:10	4:10	5:10	6:10
>	×	×	×	×	×	×	×	×	×	○	×	×	×	×	×
* j = j#6	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	7=1+6	↓	↓	↓	↓
* i = 1	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	i = 1	↓	↓	↓	↓
Note	*:conditional, d=1,0, v=1,0, ○:approve, ×:disapprove														

Table 10 The processing of the program for 10C2. (2/3)

k	2	2	3	3	3	3	3	3	3	4	4	4	4	4	4
l	9	10	4	5	6	7	8	9	10	5	6	7	8	9	10
i	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
j	7	7	7	7	7	13	13	13	13	13	13	13	13	13	13
Top:	16th	17th	18th	19th	20th	21st	22nd	23rd	24th	25th	26th	27th	28th	29th	30th
Cells(i,j)=k	(6,7)=2	(7,7)=2	(8,7)=3	(9,7)=3	(10,7)=3	(1,13)=3	(2,13)=3	(3,13)=3	(4,13)=3	(5,13)=4	(6,13)=4	(7,13)=4	(8,13)=4	(9,13)=4	(10,13)=4
k : d-1	2:10-1	2:10-1	3:10-1	3:10-1	3:10-1	3:10-1	3:10-1	3:10-1	3:10-1	4:10-1	4:10-1	4:10-1	4:10-1	4:10-1	4:10-1
≤	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
j = j#1	8=7+1	8=7+1	8=7+1	8=7+1	8=7+1	14=13+1	14=13+1	14=13+1	14=13+1	14=13+1	14=13+1	14=13+1	14=13+1	14=13+1	14=13+1
Cells(i,j)=l	(6,8)=9	(7,8)=10	(8,8)=4	(9,8)=5	(10,8)=6	(1,14)=7	(2,14)=8	(3,14)=9	(4,14)=10	(5,14)=5	(6,14)=6	(7,14)=7	(8,14)=8	(9,14)=9	(10,14)=10
l = l#1	10=9+1	11=10+1	5=4+1	6=5+1	7=6+1	8=7+1	9=8+1	10=9+1	11=10+1	6=5+1	7=6+1	8=7+1	9=8+1	10=9+1	11=10+1
l : d	10:10	11:10	5:10	6:10	7:10	8:10	9:10	10:10	11:10	6:10	7:10	8:10	9:10	10:10	11:10
>	×	×	×	×	×	×	×	×	×	○	×	×	×	×	○
* k = k+1	↓	3=2+1	↓	↓	↓	↓	↓	↓	↓	4=3+1	↓	↓	↓	↓	5=4+1
* l = k+1	↓	4=3+1	↓	↓	↓	↓	↓	↓	↓	5=4+1	↓	↓	↓	↓	6=5+1
k : d-1	2:10-1	3:10-1	3:10-1	3:10-1	3:10-1	3:10-1	3:10-1	3:10-1	3:10-1	4:10-1	4:10-1	4:10-1	4:10-1	4:10-1	5:10-1
≤	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
i = i#1	7=6+1	8=7+1	9=8+1	10=9+1	11=10+1	2=1+1	3=2+1	4=3+1	5=4+1	6=5+1	7=6+1	8=7+1	9=8+1	10=9+1	11=10+1
j = j-1	7=8-1	7=8-1	7=8-1	7=8-1	7=8-1	13=14-1	13=14-1	13=14-1	13=14-1	13=14-1	13=14-1	13=14-1	13=14-1	13=14-1	13=14-1
i : v	7:10	8:10	9:10	10:10	11:10	2:10	3:10	4:10	5:10	6:10	7:10	8:10	9:10	10:10	11:10
>	×	×	×	×	○	×	×	×	×	×	×	×	×	×	○
* j = j#6	↓	↓	↓	↓	↓	13=7+6	↓	↓	↓	↓	↓	↓	↓	↓	19=13+6
* i = 1	↓	↓	↓	↓	↓	i = 1	↓	↓	↓	↓	↓	↓	↓	↓	i = 1
Note	*:conditional, d=1,0, v=1,0, ○:approve, ×:disapprove														

Figure 12 にあるように、AO2セルからAO13セルに関数「=FREQUENCY((E1:E1000000,K1:K1000000,Q1:Q1000000,W1:W1000000,AC1:AC1000000,AI1:AI1000000),AN2:AN13)」を入力する。複数のリストを「,」で併記し、対象範囲全体を()で括る。「AN2:AN13」は前述同様、階級の区切りとなる数値を示している。この

Table 10 The processing of the program for 10C2. (3/3)

k	5	5	5	5	5	6	6	6	6	7	7	7	8	8	9	10
l	6	7	8	9	10	7	8	9	10	8	9	10	8	10	11	12
i	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6
j	19	19	19	19	19	19	19	19	19	25	25	25	25	25	25	25
Top:	31st	32nd	33rd	34th	35th	36th	37th	38th	39th	40th	41st	42nd	43rd	44th	45th	46th
Cells(i,j)=k	(1,19)=5	(2,19)=5	(3,19)=5	(4,19)=5	(5,19)=5	(6,19)=6	(7,19)=6	(8,19)=6	(9,19)=6	(10,19)=7	(1,25)=7	(2,25)=7	(3,25)=8	(4,25)=8	(5,25)=9	(6,25)=10
k : d-1	5:10-1	5:10-1	5:10-1	5:10-1	5:10-1	6:10-1	6:10-1	6:10-1	6:10-1	7:10-1	7:10-1	7:10-1	8:10-1	8:10-1	9:10-1	10:10-1
≤	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	×
j = jt1	20=19+1	20=19+1	20=19+1	20=19+1	20=19+1	20=19+1	20=19+1	20=19+1	20=19+1	20=19+1	26=25+1	26=25+1	26=25+1	26=25+1	26=25+1	
Cells(i,j)=l	(1,20)=6	(2,20)=7	(3,20)=8	(4,20)=9	(5,20)=10	(6,20)=7	(7,20)=8	(8,20)=9	(9,20)=10	(10,20)=8	(1,26)=9	(2,26)=10	(3,26)=9	(4,26)=10	(5,26)=11	
l = lt1	7=6+1	8=7+1	9=8+1	10=9+1	11=10+1	8=7+1	9=8+1	10=9+1	11=10+1	9=8+1	10=9+1	11=10+1	10=9+1	11=10+1	12=11+1	
l : d	7:10	8:10	9:10	10:10	11:10	8:10	9:10	10:10	11:10	9:10	10:10	11:10	10:10	11:10	12:10	
>	×	×	×	×	○	×	×	×	○	×	×	○	×	○	○	
* k = k+1	↓	↓	↓	↓	6=5+1	↓	↓	↓	7=6+1	↓	↓	8=7+1	↓	9=8+1	10=9+1	
* l = k+1	↓	↓	↓	↓	7=6+1	↓	↓	↓	8=7+1	↓	↓	9=8+1	↓	10=9+1	11=10+1	
k : d-1	5:10-1	5:10-1	5:10-1	5:10-1	6:10-1	6:10-1	6:10-1	6:10-1	7:10-1	7:10-1	7:10-1	8:10-1	8:10-1	9:10-1	10:10-1	
≤	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	×
j = jt1	2=1+1	3=2+1	4=3+1	5=4+1	6=5+1	7=6+1	8=7+1	9=8+1	10=9+1	11=10+1	2=1+1	3=2+1	4=3+1	5=4+1	6=5+1	
j : jt1	19=20-1	19=20-1	19=20-1	19=20-1	19=20-1	19=20-1	19=20-1	19=20-1	19=20-1	19=20-1	25=26-1	25=26-1	25=26-1	25=26-1	25=26-1	
i : v	2:10	3:10	4:10	5:10	6:10	7:10	8:10	9:10	10:10	11:10	2:10	3:10	4:10	5:10	6:10	
>	×	×	×	×	×	×	×	×	×	○	×	×	×	×	×	
* j = jt6	↓	↓	↓	↓	↓	↓	↓	↓	↓	25=19+6	↓	↓	↓	↓	↓	
* i = 1	↓	↓	↓	↓	↓	↓	↓	↓	↓	i=1	↓	↓	↓	↓	↓	
Note	* : conditional, d=10, v=10, ○ : approve, × : disapprove															

Table 11 Example of combination 2 numbers from 3276 numbers. (3276C2 sheet) (1/2)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	1	2	-1.66667	-1.5	-0.16667	322	408	0	-0.83333	0.833333			682	1948	-1	-0.5
2	1	3	1.66667	-1.5	-0.16667	322	409	0	-0.83333	0.833333			682	1949	-1	-0.33333
3	1	4	-1.66667	-1.5	-0.16667	322	410	0	-0.83333	0.833333			682	1950	-1	-0.33333
4	1	5	1.66667	-1.33333	-0.33333	322	411	0	-0.83333	0.833333			682	1951	-1	-0.33333
5	1	6	-1.66667	-1.33333	-0.33333	322	412	0	-0.83333	0.833333			682	1952	-1	-0.33333
6	1	7	1.66667	-1.33333	-0.33333	322	413	0	-0.66667	0.666667			682	1953	-1	-0.33333
7	1	8	-1.66667	-1.33333	-0.33333	322	414	0	-0.66667	0.666667			682	1954	-1	-0.33333
8	1	9	1.66667	-1.33333	-0.33333	322	415	0	-0.66667	0.666667			682	1955	-1	-0.16667
9	1	10	-1.66667	-1.16667	-0.5	322	416	0	-0.66667	0.666667			682	1956	-1	-0.16667
10	1	11	1.66667	-1.16667	-0.5	322	417	0	-0.66667	0.666667			682	1957	-1	-0.16667
11	1	12	-1.66667	-1.16667	-0.5	322	418	0	-0.5	0.5			682	1958	-1	-0.16667
12	1	13	1.66667	-1.16667	-0.5	322	419	0	-0.5	0.5			682	1959	-1	-0.16667
13	1	14	-1.66667	-1.16667	-0.5	322	420	0	-0.5	0.5			682	1960	-1	0
14	1	15	1.66667	-1.16667	-0.5	322	421	0	-0.33333	0.333333			682	1961	-1	0
15	1	16	-1.66667	-1	-0.66667	322	422	0	-0.33333	0.333333			682	1962	-1	0
16	1	17	1.66667	-1	-0.66667	322	423	0	-0.16667	0.166667			682	1963	-1	0.166667
17	1	18	-1.66667	-1	-0.66667	322	424	0	-1	1			682	1964	-1	0.166667
18	1	19	1.66667	-1	-0.66667	322	425	0	-1	1			682	1965	-1	0.333333
19	1	20	-1.66667	-1	-0.66667	322	426	0	-1	1			682	1966	-1	-0.5
20	1	21	1.66667	-0.83333	-0.83333	322	427	0	-1	1			682	1967	-1	-0.33333
21	1	22	-1.66667	-0.83333	-0.83333	322	428	0	-1	1			682	1968	-1	-0.33333

Table 11 Example of combination 2 numbers from 3276 numbers. (3276C2 sheet) (2/2)

	Q	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
999989	-0.5	0.666667	-1.16667	1624	2541	-0.16667	0.5	-0.66667			2422	3046	0.5	1	-0.5
999990	-0.5	0.833333	-1.33333	1624	2542	-0.16667	0.166667	-0.33333			2422	3047	0.5	0.666667	-0.16667
999991	-0.5	0.666667	-1.16667	1624	2543	-0.16667	0.166667	-0.33333			2422	3048	0.5	0.833333	-0.33333
999992	-0.5	0.666667	-1.16667	1624	2544	-0.16667	0.166667	-0.33333			2422	3049	0.5	0.833333	-0.33333
999993	-0.5	0.833333	-1.33333	1624	2545	-0.16667	0.166667	-0.33333			2422	3050	0.5	1	-0.5
999994	-0.5	0.833333	-1.33333	1624	2546	-0.16667	0.333333	-0.5			2422	3051	0.5	0.833333	-0.33333
999995	-0.5	1	-1.5	1624	2547	-0.16667	0.333333	-0.5			2422	3052	0.5	0.833333	-0.33333
999996	-0.5	0.666667	-1.16667	1624	2548	-0.16667	0.333333	-0.5			2422	3053	0.5	1	-0.5
999997	-0.5	0.833333	-1.33333	1624	2549	-0.16667	0.5	-0.66667			2422	3054	0.5	1	-0.5
999998	-0.5	0.833333	-1.33333	1624	2550	-0.16667	0.5	-0.66667			2422	3055	0.5	1.166667	-0.66667
999999	-0.5	1	-1.5	1624	2551	-0.16667	0.666667	-0.83333			2422	3056	0.5	1.166667	-0.66667
1000000	-0.5	0.833333	-1.33333	1624	2552	-0.16667	0.166667	-0.33333			2422	3057	0.5	0.333333	0.166667

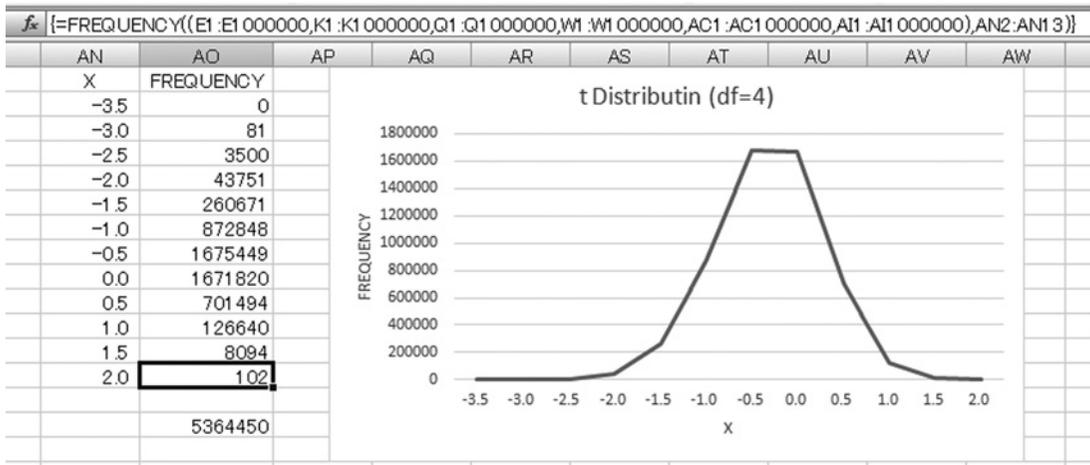


Figure 12 Frequency distribution table and t distribution curve.

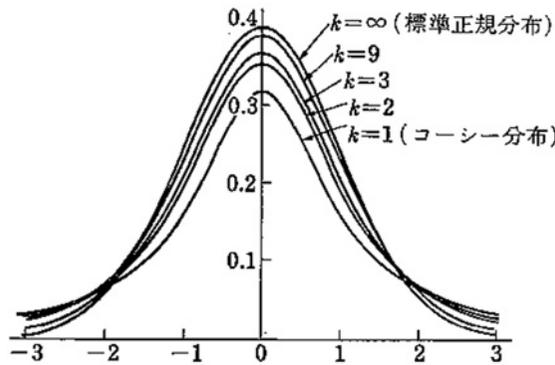


Figure 13 Probability density curve of t-distribution in various degree of freedom. (Suzuki, 1977)

関数を実行すると、AO2セルからAO13セルに度数分布が得られる。これをグラフに示したのが、FIGURE 12のグラフであり、 t 分布を示している。ここでは、標本サイズ3の2つの平均の差を問題にしているので、一つの平均を得るのに自由に変化できるデータは2つある。それゆえ、2つの平均値の場合、自由度は $2+2=4$ となる。このようにして、平均値の差の t 分布が得られることになる。自由度に応じて t 分布の確率密度曲線がFigure 13のようになる。自由度が30以上に大きくなると t 分布は標準正規分布となる。

このようにして、正規分布する同一母集団から標本を取り出し、得られた標本の平均値も t 分布をしているが、標本平均の差もまた t 分布していることがわかる。Figure 14に示すように、データとして得られた標本平均値の差が t 分布の中でどこに位置するかを見ることによって、その平均値の差が有意なものか否かを判定することができる。有意水準が5%とすると、両側検定の場合は、標本平均が(a)の棄却域(両端の斜線部2.5%)に入る時に、片側検定の場合は、標本平均が(b)の棄却域(斜線部5%)に入る時に、極めて起こりにくいことが起こったと判断する。普通ならば起こりにくいことが起こった原因がどこにあるのかと考えた場合、その原因として、得られたデータが同一母集団から抽出されたものであるというそもその前提、すな

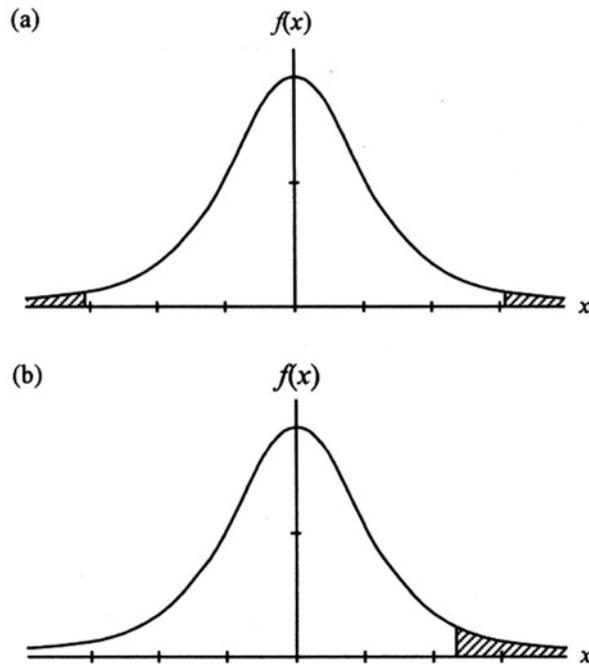


Figure 14 (a) Two-sided test and (b) One-sided test in t-distribution. Shaded area is the rejection region. (Murakami, 2002)

わち帰無仮説が間違っていたとしてこれを否定して、得られたデータは異なる母集団から得られた蓋然性が高いと判定し、有意差が現れたと考える。これが t 検定の原理である。

以上のように、本稿では、Excel の VBA を利用して、 t 分布の導出過程のシミュレーションを行ない、 t 検定の原理をたどった。標準正規分布している母集団を作成し、その中から標本サイズが 3 からなる標本を抽出し、その平均値を求めた。そして、得られた標本平均値の差の分布を求め、 t 分布が得られることを示した。

t 検定は、正規分布をしている母集団から取り出された 2 つの標本集団の平均値の差が t 分布をしていることを利用して、問題となる平均値の差が統計的に意味のあるものであるか否かを判定しようというものであることを示した。

注

- 1) OS は Windows 8.1 で、ソフトは Microsoft Excel 2013 を用いた。CPU は Intel Core i7 3.10GHz で RAM は 8.00GB を用いた。
- 2) VB (Visual Basic) については、門田 (1998, 1999, 1999b, 1999c, 2002, 2006) を参照、正規分布については、門田 (2013) を、 χ^2 分布については、門田 (2014) を参照のこと。

参考文献

- 日花弘子 2011 「仕事に役立つ Excel 統計解析 第 3 版」ソフトバンククリエイティブ
IT フロンティア 2003 「Visual Basic.NET 逆引き大全 500 の極意」秀和システム

- 岩原信九郎 1970 「教育と心理のための推計学」日本文化科学社
- 金城俊哉 2005 「Visual Basic パーフェクトマスター」秀和システム
- 門田幸太郎 1998 「VISUAL BASIC のプログラミング法1——その特徴とプログラミングの基礎——」立命館産業社会論集 第34巻 第3号 pp.119-136
- 門田幸太郎 1999 「VISUAL BASIC のプログラミング法2——配列データの操作——」立命館産業社会論集 第34巻 第4号 pp.167-187
- 門田幸太郎 1999b 「VISUAL BASIC のプログラミング法3——ファイルの操作：読み込みと表示の基礎——」立命館産業社会論集第35巻 第1号 pp.1-13
- 門田幸太郎 1999c 「VISUAL BASIC のプログラミング法4——ファイルの操作：読み込みと表示の応用——」立命館産業社会論集第35巻 第2号 pp.125-141
- 門田幸太郎 2002 「VISUAL BASIC による応答的プログラム——ユーザの成績に応じた問題提示法——」立命館産業社会論集 第38巻 第2号 pp.1-17
- 門田幸太郎 2006 「Visual Basic.NET による応答的学習プログラム——英語構文の学習に向けて——」立命館産業社会論集 第42巻 第1号 pp.161-181
- 門田幸太郎 2013 「Excel によるシミュレーションを用いた正規分布表の詳細化と Visual Basic による累積確率の検索方法」立命館産業社会論集 第49巻 第2号 pp.1-17
- 門田幸太郎 2014 「 χ^2 分布理解のための Excel によるシミュレーション—— χ^2 値の定義と分布図との関連性について——」立命館産業社会論集 第49巻 第4号
- 守谷栄一 1987 「詳解演習数理統計」日本理工出版会
- 村上雅人 2002 「なるほど統計学」海鳴社
- 武藤真介 1995 「統計解析ハンドブック」朝倉書店
- 成富慶子 2007 「EXCEL 関数辞典」秀和システム
- Norma Gilbert 1981 “Statistics second edition” Saunders college publishing
- 芝祐順・渡部洋・石塚智一 1984 「統計用語辞典」新曜社
- 鈴木義一郎 1977 「統計解析法の原理」朝倉書店
- 高木貞治 1983 「解析概論」岩波書店
- 竹村彰通 2000 「統計 第2版」共立出版
- 常見美保 2007 「EXCELVBA 辞典」秀和システム
- 山本昌弘・重定恕彦 2004 「例題でわかる Visual Basic.NET」東京電機大学出版局
- 和田秀三 1990 「基本演習確率統計」サイエンス社

Simulation Using Excel for an Understanding of t-Distribution : For the Derivation of t-Distribution Curves

MONDEN Kotaro ⁱ

Abstract : The purpose of this paper is to assist beginners in understanding intuitively both the process of making t-distribution curve and principle of t-test using Excel simulations.

First, A way to create a population of canonical normal distribution is described. Then every mean of samples consisting of 3 elements is calculated. The t-distribution curve of the difference among the means is obtained. The distribution curve of the difference among the means is obtained using frequency function in Excel.

A t-test is based on the fact that the difference between the mean values of two groups taken from the same population which has a normal distribution. The t-test showed the difference of the average value in question to be statistically meaningful or not.

Keywords : Excel, simulation, t-distribution, canonical normal distribution, NORMDIST Function, FREQUENCY Function, INDIRECT Function

ⁱ Professor, Faculty of Social Sciences, Ritsumeikan University

門田幸太郎教授 略歴と業績

I. 略 歴

- 1949年10月 大阪府に生まれる
1974年3月 神戸大学教育学部教育心理学科卒業
1977年3月 京都大学大学院教育学研究科修士課程修了
1980年3月 京都大学大学院教育学研究科博士課程満期退学
1981年4月 立命館大学文学部助教授
1994年4月 立命館大学産業社会学部助教授
1996年4月 立命館大学産業社会学部教授
2015年3月 学校法人立命館定年退職
2015年4月 立命館大学特別任用教授，名誉教授

(学内役職歴)

- 1987年4月～1988年3月 文学部学生主事

II. 専門分野

担当科目 社会心理学，情報処理統計学

研究課題 社会的認知，データ解析

学 位 教育学修士（教育方法学，京都大学，1977年4月）

所属学会 日本心理学会，日本教育心理学会，日本社会心理学会，日本行動計量学会

III. 研究業績

著 書

1. (共著)『資料で語る青年心理学』（「第7章 社会化」，斉藤稔正・井上公大編，昭和堂，1983年4月）103-116頁
2. (共著)『教育心理学』（「第5章 学級経営の心理」，富本佳郎・古厩勝彦編，福村出版，1983年11月）94-102頁
3. (共著)『教育心理学の展開』（「教師・生徒関係」，梅本克夫編，新曜社，1985年4月）216-217頁
4. (共著)『児童期の間人関係』（「第7章 児童をとりまく人間関係の問題—現代の社会病理現象—」，小石寛文編，培風館，1995年9月）118-141頁
5. (共著)『〈方法〉としての人間と文化』（「社会的比較」，佐藤嘉一編，ミネルヴァ書房，2004年）76-78頁

論 文

1. (単著)「協和・不協和条件における態度変化と態度再生」(『心理学研究』第51巻第3号, 日本心理学会, 1980年8月) 128-135頁
2. (共著)“EXPERIMENTAL STUDY OF AGGRESSION AND CATHARSIS IN JAPANESE”(Perceptual and Motor Skills Vol.58, Perceptual and Motor Skills, 1985年3月) pp.207-212.
3. (共著)「社会システムのシミュレーション—SIMSOC の概要と試験的实施1—」(『立命館文学』第475-477号, 1985年3月) 1-64頁
4. (共著)「SIMSOC 進行係用マニュアルⅠ」(『立命館文学』第478-480号, 1985年6月) 55-96頁
5. (共著)「SIMSOC 進行係用マニュアルⅡ」(『立命館文学』第483-484号, 1985年10月) 51-80頁
6. (単著)「C言語でのプログラミングによるデータ解析法の理解—標本特性値と母数との関係—」(『立命館産業社会論集』第30巻第3号, 1994年12月) 123-170頁
7. (単著)「VISUAL BASIC のプログラミング法1—その特徴とプログラミングの基礎—」(『立命館産業社会論集』第34巻第3号, 1998年12月) 119-136頁
8. (単著)「VISUAL BASIC のプログラミング法2—配列データの操作—」(『立命館産業社会論集』第34巻第4号, 1999年3月) 167-187頁
9. (単著)「VISUAL BASIC のプログラミング法3—ファイルの操作;読み込みと表示の基礎—」(『立命館産業社会論集』第35巻第1号, 1999年7月) 1-13頁
10. (単著)「VISUAL BASIC のプログラミング法4—ファイルの操作;読み込みと表示の応用—」(『立命館産業社会論集』第35巻第2号, 1999年9月) 125-141頁
11. (共著)「子どもの存在と地域の教育力—地域差からの考察—」(『教育心理学フォーラム・レポート FR-2002-003』, 日本教育心理学会, 2002年7月) 1-19頁
12. (単著)「VISUAL BASIC による応答的プログラミング—ユーザの成績に応じた問題提示法—」(『立命館産業社会論集』第38巻第2号, 2002年9月) 1-16頁
13. (共著)「対人認知における類似性と非類似性について」(『立命館産業社会論集』第40巻第3号, 2004年12月) 21-36頁
14. (単著)「Visual Basic.NET による応答的学習プログラム—英語構文の学習に向けて—」(『立命館産業社会論集』第42巻第1号, 2006年6月) 161-181頁
15. (共著)「日本, 中国, 韓国における親の養育態度と子どもの学業志向性及び向社会的仲間志向性との関係」(『教育心理学フォーラム・レポート FR-2012-02』, 日本教育心理学, 2013年1月) 1-6頁
16. (単著)「Excel によるシミュレーションを用いた正規分布表の詳細化と Visual Basic による累積確率の検索方法」(『立命館産業社会論集』第48巻第4号, 2013年3月) 1-12頁
17. (単著)「 χ^2 分布理解のための Excel によるシミュレーション— χ^2 値の定義と分布図との関連性について—」(『立命館産業社会論集』第49巻第4号, 2014年3月) 1-14頁

その他: 書評

(単著)「ロジャー・C. マンネル, ダグラス・A. クリーバー著, 速水敏彦(監訳)『レジャーの社会心理学』2004年, 世界思想社」(『社会心理学研究』第20巻第3号, 日本社会心理学会, 2005年3月) 253-254頁

IV. 社会活動

2006年4月～2007年3月 京都地方裁判所委員会委員