

## 研究ノート

# 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

門田 幸太郎\*

本稿の目的は次の3つである。第1に、Excelを用いて正規分布曲線を作成することを通して、確率密度関数と累積確率との関係を示すこと。第2に、区分求積法の幅を0.001とすることによって、小数点以下4桁の正規確率分布表を作成すること。第3に、 $z$ 値に対応する累積確率を検索する方法として Visual Basic の利用方法を示すことである。

キーワード：正規分布，確率密度関数，累積確率，区分求積法，Visual Basic，Excel VBA によるシミュレーション

## 1. 正規分布

正規分布 (normal distribution) は統計学の基礎である。統計的検定のもとになる  $t$  分布, F 分布,  $\chi^2$  分布などの理解も、正規分布の理解があって初めて可能となる。ここでは、Excel を用いて、正規分布関数の累積確率値を求め、正規分布表を作成する方法について説明する。この方法を理解することを通して、正規確率密度関数の性質を理解するとともに、より詳細化された正規分布表を作成することができるようになる。現在、広く用いられている正規分布表は小数点以下3桁のものが一般的である。しかし、より詳細な値を求めることが必要となる場合がありうる。そのような場合でも、自分自身で、正規分布表を作成することができれば、必要に応じて、より詳細な数値を求めることができるようになる。また、累積確率を求めるために Visual Basic を利用する方法について説明する。

### 1-1. 確率密度関数

平均  $\mu$ 、標準偏差  $\sigma$  の正規分布の確率密度関数はガウス関数の一種で、

---

\* 立命館大学産業社会学部教授

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

で求められる。これは、 $\mu$  を中心として、分散が  $\sigma^2$  の凸のベル型の曲線を表している。 $e$  は自然対数を示す。

ここで、

$$z = \frac{x-\mu}{\sigma} \text{ と変換してみると、この曲線は中心が } 0 \text{ となり、分散が } 1 \text{ となる。この } z \text{ 変換を用いる}$$

ことにより、 $\mu = 0$ 、 $\sigma = 1$  を代入すると、上式は

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

となる。ここで、 $x$  を  $z$  に書き換えると、

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

となる。これを正規確率密度関数 (standard normal distribution function) という。与えられたデータを標準化することにより、 $f(z)$  式を適用すると確率密度を求めることができる。

確率密度関数の導出は 2 項分布  $nCrp^x(1-p)^{n-x}$  の階乗  $nCr = \frac{n!}{r!(n-r)!}$  に Stirling の公式<sup>1)</sup>

$n! \cong \sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n}$  をあてはめることによって得られる。2 項分布は数学の基礎的な知識と言えるが、Stirling の公式となると初学者には理論的に理解が困難となる。そこで、区分求積法によるシミュレーションにより正規分布の性質を直観的に理解することが、理論的理解へのひとつの手掛かりになりうる。もちろん、直観的理解に留まっているだけでは不十分である。しかし、まずは理論的に理解しようという内容がいかなる事柄——数学的現象——を取り扱っているのかをイメージできることが理解へのひとつのアプローチ方法と言える。

## 1-2. 自然対数 $e$ のシミュレーション

最初に、確率密度関数で用いられる自然対数について考えてみる。自然対数  $e$  は、

$$\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e$$

によって定義されるものである。これは  $x$  を限りなく大きくした場合に  $\left(1 + \frac{1}{x}\right)^x$  が一定の値  $e$  に収束することを意味している。 $e$  が存在することは 2 項分布の展開を用いて証明することができる。しかし、ここでは、より直観的に理解できるように、Excel によるシミュレーションを利用して、 $x$  が増大するにつれて、 $e$  が収束するという (数学的) 現象を示すことにする。

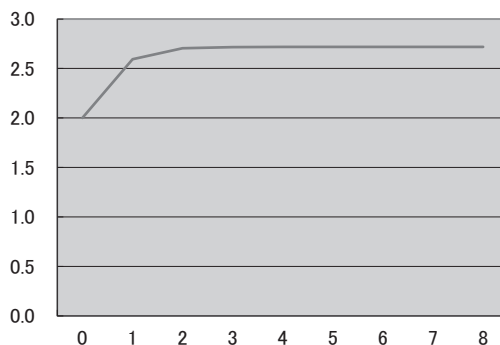
手順としては、まず、Excel の Sheet の B1 セルに  $x$  の初期値として 1 を入力し、C1 セルには  $1 + \frac{1}{x}$

Table 1. シミュレーションによる自然対数  $e$  の値

	A	B	C	D
1		$x$	$1 + \frac{1}{x}$	$\left(1 + \frac{1}{x}\right)^x$
2	0	1	2.0000000000	2.0000000000
3	1	10	1.1000000000	2.5937424601
4	2	100	1.0100000000	2.7048138294
5	3	1000	1.0010000000	2.7169239322
6	4	10000	1.0001000000	2.7181459268
7	5	100000	1.0000100000	2.7182682372
8	6	1000000	1.0000010000	2.7182804692
9	7	10000000	1.0000001000	2.7182816940
10	8	100000000	1.0000000100	2.7182817864

を求める関数として「 $=1+1/B1$ 」を入力し、D1セルには $\left(1+\frac{1}{x}\right)^x$ を求める関数として「 $=POWER(C1,B1)$ 」を入力する。POWER関数はべき乗を求めるものであり、C1セルには、べき乗を求める数値である底を指定し、B1セルには何乗するかを示す指数を指定する。次にB2セルに「 $=B1*10$ 」を入力する。これは $x$ の増加率を10倍とすることを意味する。B2セルに入力された関数をB9セルまでコピーする。これにより、 $x$ の値が10から $10^8$ まで増大することになる。C1セルとD1セルの関数をC9セルとD9セルまでコピーする。これにより、 $x$ の増大に応じて $1+\frac{1}{x}$ がC列に、 $\left(1+\frac{1}{x}\right)^x$ がD列に表示されたTable 1を得ることができる。ここから、 $x$ の値が等比級数的に増大すれば、次第に収束することが直観的に理解できるようになる。ExcelのEXP関数を用いるとEXP(1)で $e$ の値が2.718281828であることがわかる。Table 1から $x$ の値が $10^2$ の時、近似値は小数点以下第2位まで近似しており、 $10^5$ の時、第4位まで、 $10^8$ の時、第6位まで近似していることがわかる。

Figure 1はD1セルからD9セルに得られた近似値を示したものである。縦軸は近似値を、横軸は

Figure 1. 自然対数  $e$  の近似値の図示

指数を示す数値である。この数値はA1セルからA9セルまでに納められている。これは求めるべき乗の指数を示している。たとえば、横軸の4は $x$ の値が $10^4$ であることを意味している。D列を系列としてグラフ表示してみると、 $e$ の値が収束している様子が見られる。

## 2. 正規分布表の作成方法

### 2-1. 区分求積法による累積確率

一般に用いられている正規分布表には2種類ある。正規分布曲線と $z$ 軸に挟まれた面積のうち、 $Z$ 値が0から特定の $Z$ 値までの積分 $\int_0^z f(t)dt$ に相当する面積を示すタイプのもとの、特定の $Z$ 値以上の積分 $\int_z^{+\infty} f(t)dt$ に相当する面積を示すタイプのものがある。本稿では、前者の表示方法をとる。これは、確率密度関数 $f(t)$ を0から $z$ までを積分したものであり、 $\int_0^z f(t)dt$ という数式で求められる面積を表している。Figure 2の横線の網掛けになった部分に相当する。この面積を、区分求積法により求める。正規分布曲線と $z$ 軸に挟まれた幅0.001の面積を、その幅の中間点と分布曲線との交点を高さとした幅0.001の長方形でシミュレートする。Figure 2の「相殺」にあるように、長方形からはみ出た黒く塗りつぶされた三角形と長方形から欠けた斜線が付けられた三角形とが相殺すると考えられる。このようにして、 $\int_0^z f(t)dt$ の積分の近似値として長方形の面積の合計されたものを利用することができる。確率密度関数は、各 $z$ 値の確率そのものを示すものではなく、確率の密度を示している。確率そのものは確率密度関数を積分した結果 i.e. 面積によって示される。ゆえ

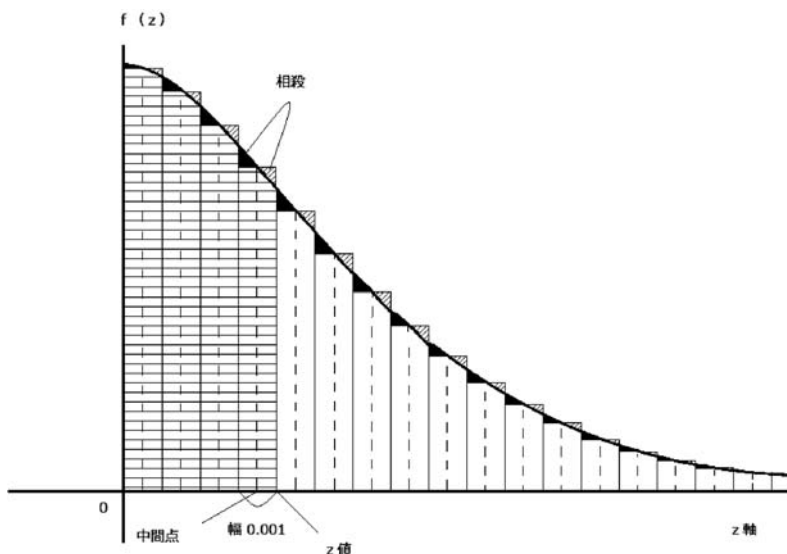


Figure 2. 区分求積法による累積確率の模式図

に、中間点と曲線との交点の距離（高さ）と幅との積を求めて初めて各幅での確率が求められる。これらの各長方形の面積を足し合わせたものが累積確率となる。

## 2-2. Excel による正規分布表の作成

$\frac{1}{\sqrt{2\pi}}$  を係数とし、 $e$  を底、 $-\frac{z^2}{2}$  を指数とする確率密度関数  $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$  を求める。

Excel の sheet を開き、A1 セルに、初期値として 0 を入れる。次に、A2 セルに、0.001 単位の等差数列を作るために、関数「=a1+0.001」を入れる。さらに、B2 セルに、 $z$  軸の区間の中間点を求めるため、関数「=AVERAGE(A1,A2)」を入れる。ここでは、A1 セルの 0.000 と A2 セルの 0.001

との中間点の 0.0005 を求めることになる。C2 セルには、確率密度関数  $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$  の  $z$  値として

B2 セルの値を入れて求められた結果である確率密度の値を示す。D2 セルには、この確率密度と区間の幅である 0.001 との積を求める関数「=C2\*0.001」を入れる。これは 0.001 幅の区分での確率を示している。最後に、E2 に累積確率を求める関数「=E1+D2」を入れる。これにより、0 からこの  $z$  値までの確率密度関数の曲線と  $Z$  軸で挟まれた面積を示している。次に、A2 セルから E2 セルまでをフォーカスし、A3 セルから E5454 セルまでドラッグ、リリースしてコピーする。ちなみに、

正規分布関数 (normal distribution function) は、 $F(z) = \int_{-\infty}^z f(t)dt$  として定義される。この値は、E 列

で求められた累積確率に  $\int_{-\infty}^0 f(t)dt = 0.5$  を加えることによって得られる。E 列で求められた累積確率

は数式で表現すると  $\int_0^z f(t)dt$  となる。

次に、列ごとにセル書式を設定する。A 列をフォーカスし、プルダウンメニューからセルの書式設定を選び、表示形式のタグにある分類の中から数値を指定し、小数点以下の桁数を 3 とし、「OK」ボタンをクリックする。B 列以下も同様に設定するが、小数点以下の桁数を 4 とし、C ~ E 列は桁数を 7 とする。C 列を系列としてグラフ表示したものが Figure 3 である。ここで縦軸は確率密度を、横軸はセルの行番号を示している。

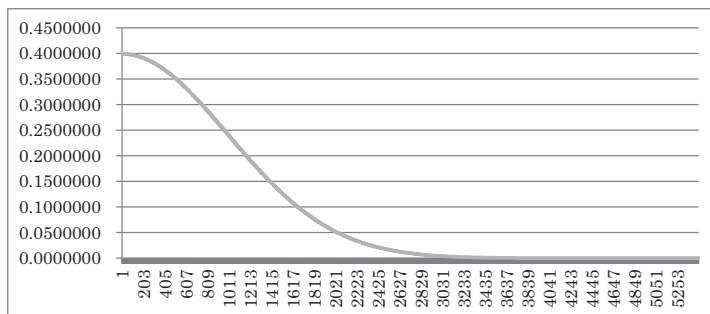


Figure 3. 正規分布曲線の図示 ( $z > 0$  の場合)

以上のような方法で得られた数表の一部を Figure 4 に示す。2580行を例に数表の数値を説明する。A2580セルには  $z$  値2.579が入力されている。B2580セルには、A列の値とひとつ上のA2579セルの2.578との中間点の値2.5785が示されている。C2580セルには確率密度0.0143606が、D2580セルには0.001幅での確率0.0000144が、E2580セルには、ひとつ上のE2579セルに納められているそれまでの累積確率0.4950313と新たに求められたD2580セルの0.0000144との合計0.4950457が新たな累積確率として示される。

一般に1%の有意水準で統計的検定を行う場合、基準となる  $z$  値は2.58が用いられる。しかし、Figure 4 のE列を見ると、両側で1%の検定を行う場合、分布の両端に  $0.01/2=0.005$  だけを残した累積確率が0.4950000 ( $=0.5-0.01/2$ ) となる  $z$  値を求めることになる。小数点以下3位までの  $z$  値を利用した場合、2.576が  $z$  値として最も近い基準値であることが示されている。

	A	B	C	D	E
1952	1.951	1.9505	0.0595366	0.0000595	0.4744715
1953	1.952	1.9515	0.0594206	0.0000594	0.4745309
1954	1.953	1.9525	0.0593047	0.0000593	0.4745902
1955	1.954	1.9535	0.0591890	0.0000592	0.4746494
1956	1.955	1.9545	0.0590735	0.0000591	0.4747085
1957	1.956	1.9555	0.0589581	0.0000590	0.4747674
1958	1.957	1.9565	0.0588429	0.0000588	0.4748263
1959	1.958	1.9575	0.0587278	0.0000587	0.4748850
1960	1.959	1.9585	0.0586129	0.0000586	0.4749436
1961	1.960	1.9595	0.0584982	0.0000585	0.4750021
1962	1.961	1.9605	0.0583837	0.0000584	0.4750605
1963	1.962	1.9615	0.0582693	0.0000583	0.4751188
1964	1.963	1.9625	0.0581551	0.0000582	0.4751769
1965	1.964	1.9635	0.0580411	0.0000580	0.4752350
1966	1.965	1.9645	0.0579272	0.0000579	0.4752929
1967	1.966	1.9655	0.0578135	0.0000578	0.4753507
1968	1.967	1.9665	0.0576999	0.0000577	0.4754084
1969	1.968	1.9675	0.0575865	0.0000576	0.4754660
1970	1.969	1.9685	0.0574733	0.0000575	0.4755235
1971	1.970	1.9695	0.0573603	0.0000574	0.4755808

$z$  値 = 1.96 (有意水準 5%) 付近

	A	B	C	D	E
2572	2.571	2.5705	0.0146594	0.0000147	0.4949297
2573	2.572	2.5715	0.0146218	0.0000146	0.4949444
2574	2.573	2.5725	0.0145842	0.0000146	0.4949589
2575	2.574	2.5735	0.0145467	0.0000145	0.4949735
2576	2.575	2.5745	0.0145093	0.0000145	0.4949880
2577	2.576	2.5755	0.0144720	0.0000145	0.4950025
2578	2.577	2.5765	0.0144348	0.0000144	0.4950169
2579	2.578	2.5775	0.0143976	0.0000144	0.4950313
2580	2.579	2.5785	0.0143606	0.0000144	0.4950457
2581	2.580	2.5795	0.0143236	0.0000143	0.4950600
2582	2.581	2.5805	0.0142867	0.0000143	0.4950743
2583	2.582	2.5815	0.0142498	0.0000142	0.4950885
2584	2.583	2.5825	0.0142131	0.0000142	0.4951027
2585	2.584	2.5835	0.0141764	0.0000142	0.4951169
2586	2.585	2.5845	0.0141398	0.0000141	0.4951311
2587	2.586	2.5855	0.0141033	0.0000141	0.4951452
2588	2.587	2.5865	0.0140669	0.0000141	0.4951592
2589	2.588	2.5875	0.0140306	0.0000140	0.4951733
2590	2.589	2.5885	0.0139943	0.0000140	0.4951872
2591	2.590	2.5895	0.0139581	0.0000140	0.4952012

$z$  値 = 2.58 (有意水準 1%) 付近

Figure 4. 小数点以下3位まで求められた正規分布表の一部



### 3. Excel の VBA を利用した累積確率の表示

#### 3-1. オブジェクトの導入とそのプロパティの設定

正規分布表を作成した sheet を利用して、Excel での Visual Basic<sup>2)</sup> を利用した累積確率を表示する。Excel には Visual Basic が組み込まれている。機能は限定されているとはいえ、かなりの性能を持っている。Visual Basic を利用する場合 3 つの手続きが必要となる。第 1 に、オブジェクトの導入とそのプロパティの設定。第 2 に、オブジェクトに対応したコードの記述 (プログラミング)。第 3 に、実行と保存。

Visual Basic を利用するためには、sheet の上にある「開発」タブをクリックして、コードのエリアにある「Visual Basic」マークを選択することから始まる。ユーザフォームの挿入のアイコンをクリックし、「プロジェクト - VBAProject」の下にある「オブジェクトの表示」をクリックする。フォームデザイナーの中に UserForm1 が表示され、それと同時に、ツールボックスも表示される。それから、テキストボックスとラベル 2 つ、コマンドボタンを 2 つ取り込み、Figure 5 のように配置する。配置後、それぞれのプロパティの設定を行う。TextBox1 の Font のサイズのプロパティを 16 としておく。Label1 の Caption を「z 値から累積確率を求める。」とする。これは、プロパティボックスの中の Caption の項目を選び、その右にある Label1 を書き換えることによって行われる。Label2 についても同様に Caption を「上のテキストボックスに z 値を入力してください。」とし、CommandButton1 の Caption を「z 値から累積確率を求める。」とし、CommandButton2 の

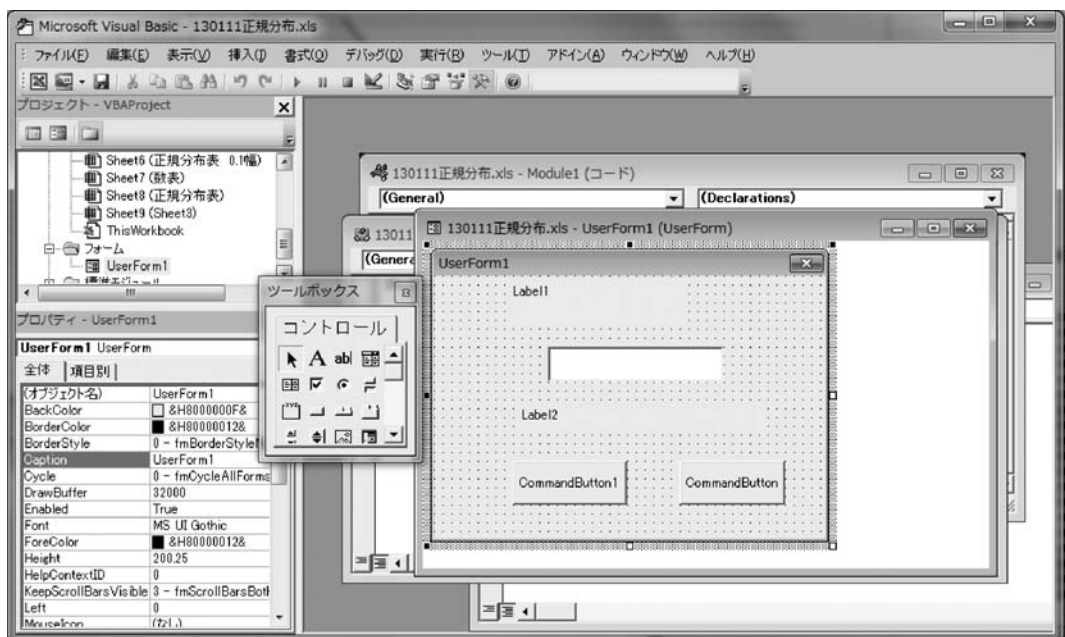


Figure 5. Excel での Visual Basic でオブジェクトを取り込んだ画面

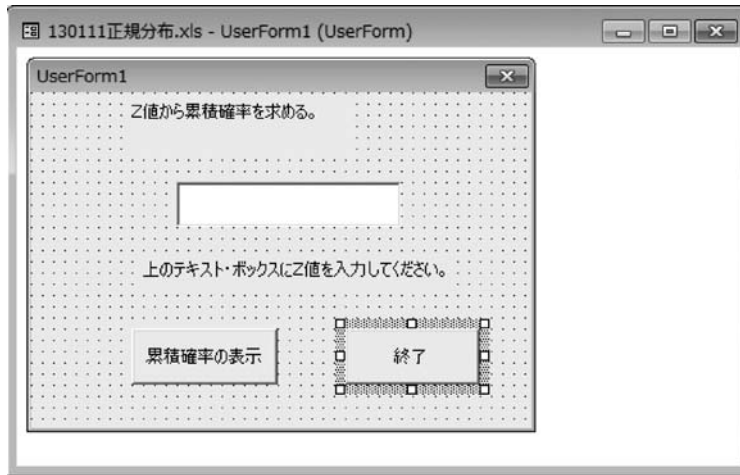


Figure 6. オブジェクトのプロパティの設定

Caption を「終了」とする。

### 3-2. コード記述

CommandButton1 をダブルクリックすると、Figure 7 のようなコード記述画面が現れる。ここに Figure 8 のようにコードを入力する。基本的な作業過程としては、先に作成した正規分布の sheet の A 列から、TextBox1 に入力された  $z$  値に相当する数値を検索し、その同じ行の E 列にある累積確率を Label1 に表示させるプログラムを作成することである。

For-Next 文はその間にあるコマンドを5454回繰り返すことを意味する。If-End If 文は If と Then との間に示された条件を満たした場合に Then 以下 End If までのコマンドを実行させるという働きをする。ここでの条件は「Cells(i,1) = CSng(TextBox1.Text)」となる。Cells(i,1) は  $i$  の値の増加に応じて、A 列の 1 から5454まで変化し、A1 セルから A5454 セルまでを検索することになる。右辺の TextBox1.Text はテキストボックスに入力された数字をテキスト型として取得したものを示す。CSng() はデータ変換の関数で、() 内に示されたテキスト型データを単精度の実数型データに変換させる。この条件文によりテキストボックスに入力された数字と等しい数値を A 列から検索することができる。

If Then コマンドの次の 2 行は、条件が満たされた場合に Label 2 のプロパティを設定するものである。「Label2.Font.Size=16」は Label 2 に表示する文字（キャプション）のフォントサイズを 16 とするものである。次の「Label2.Caption.Format(Cells(i,5), "累積確率は 0.0000000 です。")」は、Label 2 のキャプションの書式を設定している。「累積確率は」と「です。」の間にある 0.0000000 は数字ではなく、表示する書式を表している。文字内容は、sheet 内の  $i$  行 5 列のセル Cells(i,5) に含まれているものであり、その表現形式は 8 桁の小数を小数点以下 7 桁で表示し、その前後を「累積確率は」と「です。」で挟むという書式を指定している。



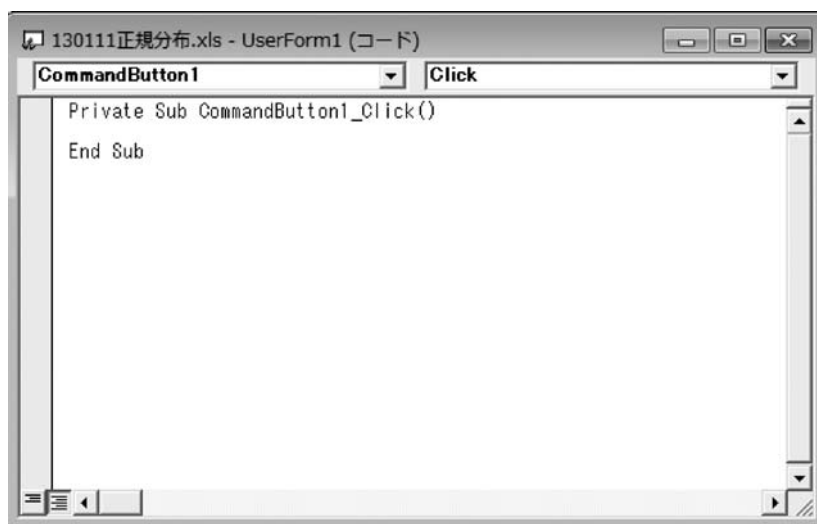


Figure 7. コード記述開始画面



Figure 8. 作成されたコード記述

CommandButton2をダブルクリックし、END コマンドを入力する。これにより、「終了」ボタンをクリックすると Visual Basic を終了することになる。

### 3-3. Visual Basic の実行と保存

コード記述が完了したならば、実行となる。このまま Visual Basic 編集画面で実行する場合は、Figure 9 に示されたように、ツールバーにある緑色の三角形の実行ボタンをクリックすることによりプログラムが実行される。Excel 画面から実行する場合は、「開発」タグをクリックし、コードのエリアにある「Visual Basic」マークをクリックすれば、Visual Basic 編集画面となる。その後の操作方法は上述の通りである。Visual Basic 編集画面から Excel 画面に移動する場合はツールバ

ーにある「Excel」アイコンをクリックすればよい。

テキストボックスに累積確率を求めるために  $z$  値を入力する。 $z$  値入力後、累積確率の表示というキャプションがついた CommandButton 1 をクリックするとラベルに入力された  $z$  値に対応する累積確率が表示される。

最後に終了ボタンをクリックすることにより、Visual Basic の実行を終了する。

このようにして作成された Visual Basic のプログラムはコード記述画面の上部にあるメニューバ

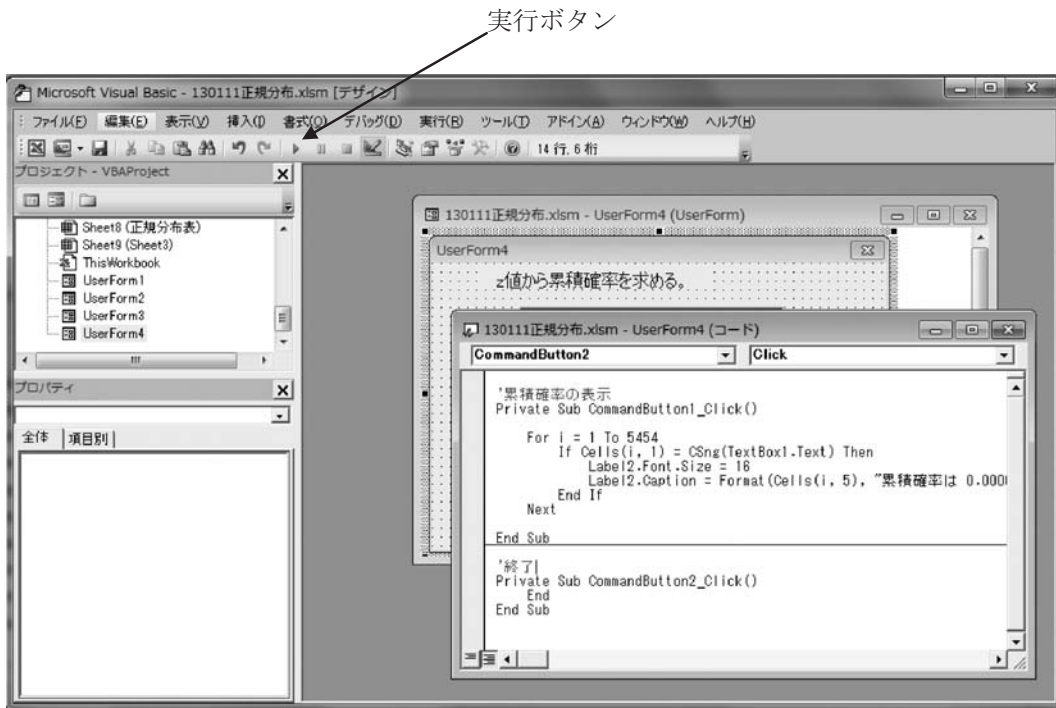


Figure 9. Visual Basic の実行



Figure 10. Visual Basic の実行例

一の「ファイル(F)」をクリックすると現れるプルダウンメニューから「保存」を指定することによって保存されることになる。

## 注

- 1) 高木貞治 (1983)
- 2) Visual Basic の利用法については門田 (1998, 1999, 1999b, 1999c, 2002, 2006) を参照のこと。

## 参考文献

- 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
- Microsoft Corporation 2002 マイクロソフト (株) 訳「Microsoft.Basic.NET」日経 BP ソフトプレス
- 河西朝雄 2003 「VB.NET 基礎学習 Bible」技術評論社
- 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——」立命館大学産業社会論集 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——」立命館大学産業社会論集 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——」立命館大学産業社会論集 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——」立命館大学産業社会論集 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——」立命館大学産業社会論集 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——」立命館大学産業社会論集 第42巻 第1号 pp. 161-181
- 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
- 芝祐順・渡部洋・石塚智一 1984 「統計用語辞典」新曜社
- 高木貞治 1983 「解析概論」岩波書店
- 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
- 若山芳三郎 2004 「学生のために Visual Basic.NET」東京電機大学出版局

## Research Note

### Refinement of Normal Distribution Table Using Simulation of Excel and Retrieval Method to Find the Cumulative Probability by Visual Basic.

MONDEN Kotaro \*

**Abstract:** This paper has the following three purposes. First, creating a normal distribution curve by Excel, it shows the relationship between the probability density function and the cumulative probability. Second, using the mensuration by division method with the width of 0.001, it is possible to create a normal probability distribution table of four decimal places. Third, using Visual Basic, shows how to find the cumulative probability corresponding to the  $z$  value.

**Keywords:** Normal Distribution, Probability Density Function, Cumulative Probability, Mensuration by Division, Visual Basic, Simulation by Excel VBA

---

\*Professor, Faculty of Social Sciences, Ritsumeikan University