

正規確率計算アルゴリズム再考

— コンピュータ支援による2項近似アルゴリズム I —

平井 安久 ・ 中村 忠*

従来とは逆に2項確率で正規確率を近似する計算法を提案する。この近似計算法は、中心極限定理、数値微分、低い次数の多項式と巨大な n に対する2項確率の直接計算法を組み合わせて構成される。これまでの他の正規確率計算法と違い、コンピュータの機能を十分に利用した方法である。数値実験により、関数項級数展開を用いた近似方法、有理関数近似を用いた近似方法、連分数展開を用いた近似方法に比べ、提案した方法は広範囲において絶対相対誤差の基準ではるかに良いことも示された。

Keywords : 2項近似, 近似法, アルゴリズム, 正規確率

1. はじめに

標準正規分布関数 $\Phi(z)$ の値(以下、正規確率という)を求めるアルゴリズムの開発は1940年代から本格的に始まった。コンピュータの進歩と共に非常に多くのアルゴリズムが提案されてきた。その手法は関数項級数(べき級数, 三角級数, 有理関数等)展開を用いた近似方法, 有理関数近似を用いた近似方法, 連分数展開(有理関数の極限值)を用いた近似方法, 正規分布に近い分布を用いた方法, ミルズ比の関数項級数表示を用いて近似方法, 定積分を直接計算する数値積分法のいずれかである。ここでは近似法に焦点を当てるので、数値積分法による方法は議論しない。各手法においては、いろいろな近似式が提案されてきた(例えば、Hastings(1955), 山内・森口・一松(1970), Abramowitz and Stegun(1972), Johnson, Kotz and Balakrishnan(1994), Patel and Read(1996))。

本論では、中心極限定理、数値微分、べき級数展開と広範囲の n に対する2項確率 $B(x; n, 0.5)$ の直接計算法を利用することにより、離散量のみを用いて正規確率 $\Phi(z)$ の値を近似する方法を提案する。従来の方法は明確な表現である数式を用いて正規確率 $\Phi(z)$ の値を近似計算する方法であることから、今日のコンピュータの強力な数値処理能力を十分に活用した方法とはいえない。また、数式による近似法の精度は狭い区間で比較的良好であるが、十分幅の広い区間では悪くなるという欠点を持つ。一方、本論で提案する方法は数式も用いるが、計算過程の一部に明確な形の数式では表せない、コンピュータでのみ実行可能な計算を含んでいる。この意味で、従来の方法とは異なり、コンピュータの機能を十分に利用した新しい正規確率 $\Phi(z)$ の近似計算法である。

第2節で、分割関数と呼ばれる関数を導入する。この関数は、その名の通り、区間を分割する役割をする。補正された正規関数、Cornish-Fisher関数、Peizer-Pratt関数は分割関数であることが示される。これらの関数は、従来の2項確率の正規近似法においては、それぞれ、補正された正規変換、Cornish-Fisher変換、

岡山大学大学院教育学研究科 自然教育学系 数学教育講座 700-8530 岡山市北区津島中3-1-1

*岡山理科大学総合情報研究科 情報科学専攻 情報理数系 700-0005 岡山市北区理大町1-1

Computation Algorithm Revisited for Normal Probability: Computer-assisted Algorithm for Binomial Approximation (1)

Yasuhisa HIRAI and Tadashi NAKAMURA*

Department of Mathematics Education, Division of Natural Science Education, Graduate School of Education, Okayama University, 3-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530

*Department of Mathematics Information Science, Okayama University of Science, 1-1 Ridaicho, Kita-ku, Okayama city 700-0005

Peizer-Pratt 変換と呼ばれている。実数 z が与えられたとき、分割関数によってある点 \hat{z} が決定される。点 \hat{z} での 5 次までの $\Phi(t)$ のテーラー展開式を使用する。テーラー展開式に現れる $\Phi(\hat{z})$ は計算不可能なので、その近似値を求めるために中村・平井(2006) が提案した 2 項確率の直接計算法とこの計算方法を修正した計算法の 2 種類を利用する。更に、1 次以上のテーラー展開式は正規確率密度関数(いわゆる、連続量) $\phi(x)$ を含んでいる。ここでは全ての量を離散量から計算することを前提としているので、この $\phi(x)$ の近似値を求めるために、4 種類の数値微分を導入する。これらを組み合わせることにより、正規確率 $\Phi(z)$ の値を離散量のみで計算する種々の近似法が導かれる。

第 3 節では、第 2 節で導出された各近似法の全体を基にし、数値実験を通じてこれらのどの近似法よりも良いことが期待される近似法の構成手続きを提案する。区間 $[-8,8]$ を長さ 2 の部分区間に分割し、提案した手続きに従って各部分区間毎に近似法を構成していく。第 3.1 節では、近似法の精度の良さ具合を調べる数値実験の環境設定や基準量を導入する。計算区間として、 $[-8,8]$ を採用する。区間 $[-8,8]$ の任意の部分区間において、2 つの近似法の組に対して良さを測る基準量を導入する。導入された基準量は計算が困難なので、 $[-8,8]$ 内に 1600 点を取り、これらの点の計算値を利用して近似法の区間における基準量の近似値を算出するアルゴリズムを提案する。第 3.2 節では、前節で導入した基準量の近似値を用いて、区間 $[-8,8]$ の任意の部分区間で近似法を構成できる。この手続きは与えられた近似法の集合に対して、この集合のどのアルゴリズムよりも良い精度が期待される新しい近似法を構成する。3 段階からなる手続きを提案する。第 3.3 節では、数値実験により前節で提案された近似法の構成手続きを長さ 2 の部分区間 $I_1 = [-8, -6)$, $I_2 = [-6, -4)$, \dots , $I_7 = [4, 6)$, $I_8 = [6, 8]$ に対して実行する。提案された近似法の精度を数値実験により求める。その結果、各区間ごとに提案する近似法が導出される。最大絶対相対誤差および計算時間の点から提案された近似法の実用性を考える。第 4 節では、提案されたアルゴリズムについて、いろいろな文献で紹介・検証されている近似法との精度比較を行う。本論では良さ具合の基準として最大絶対相対誤差を採用しているが、これまでの文献では精度は最大絶対誤差で与えられている場合が多い。そのため、ここでは 2 種類の基準で比較・検討し、提案された近似法の良さを検討する。

2. 正規確率の 2 項近似

実数 $z \in (0, \infty)$ が与えられたとき、正規確率 $\Phi(z)$ の値を 2 項確率 $B(x; n, p)$ で精度の高い近似をするために p, n, x をどのように決定すればよいのかを考える。数多くある 2 項分布 $B(n, p)$ から 1 つの正規確率 $\Phi(z)$ を近似すればよいので、2 項分布のパラメータ p を自由に選択・固定できる。正規密度関数は対称性があるので、2 項確率関数も対称性を持つように $p=0.5$ とするのは自然である。変動するパラメータは n であり、ド・モアブル＝ラプラスの中心極限定理より、 n は十分大きな自然数を選択すればよい。以下、どのようにして 2 項確率 $B(x; n, 0.5)$ で正規確率 $\Phi(z)$ の値を近似計算するかを述べよう。

定数 k_1, k_2 は非負の整数で既知とし、 $I(k_1, k_2) = \{k_1, k_1+1, \dots, n-k_2-1, n-k_2\}$ とおく。 $n \geq k_1 + k_2 + 3$ を満たす任意の整数 n に対し、関数 $d(x, n)$ ($x \in I(k_1, k_2)$) が分割関数とは次の条件が満たされる時にいう。

- (a) $d(x, n)$ は x に関して $I(k_1, k_2)$ 上で狭義単調増加関数である。
- (b) 任意の実数 $z \in (-\infty, \infty)$ に対して、ある正整数 $n_0 = n_0(z)$ が存在して、 $n > n_0$ なる限り、
 $d(\hat{x}, n) \leq z < d(\hat{x}+1, n)$ となる正整数 $\hat{x} = \hat{x}(z, n)$ ($k_1 < \hat{x} < n - k_2$) が存在する。

(c) $\lim_{n \rightarrow \infty} d(\hat{x}+i, n) = z, i = -1, 0, 1, 2 ;$

$$\lim_{n \rightarrow \infty} |B(\hat{x}+i, n, 0.5) - \Phi(d(\hat{x}, n))| = 0, i = -1, 0, 1, 2.$$

分割関数 $d(x, n)$ は、その名のとおりに、ある軸を分割する (図 2.1 参照)。

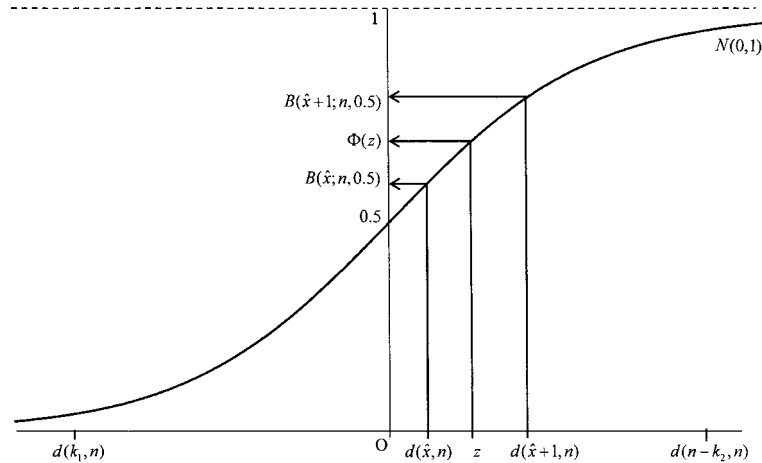


図 2.1. 分割関数 $d(x, n)$ により, z 軸の区間 $[d(k_1, n), d(n-k_2, n)]$ の分割 $[d(k_1, n), d(k_1+1, n)]$, $[d(k_1+1, n), d(k_1+2, n)]$, \dots , $[d(k_1, n), d(n-k_2, n)]$ ができる.

以下, 本論を通じて, $\hat{x} = \hat{x}(z, n)$, $z_i = d((\hat{x}+i), n)$ ($-1 \leq i \leq 2$), $u(x, n) = (2x+1-n)/\sqrt{n}$, $a(x, n) = x - (n+1)/2 + 0.01(1/(x+1) - 1/(n-x))$ とおく。分割関数の例としては,

- (i) 補正された正規関数 ($k_1 = 0$, $k_2 = 0$)

$$d(x, n) = u(x, n), \quad 0 \leq x \leq n,$$

- (ii) Cornish-Fisher 関数 ($k_1 = 0$, $k_2 = 0$)

$$d(x, n) = u(x, n) + \frac{1}{12n} (u(x, n)^3 - u(x, n)), \quad 0 \leq x \leq n,$$

- (iii) Peizer-Pratt 関数 ($k_1 = 0$, $k_2 = 0$)

$$d(x, n) = \frac{a(x, n)}{|x+0.5-n/2| \sqrt{1+6n}} \times \left[(x+0.5) \log \frac{2x+1}{n} + (n-x-0.5) \log \frac{2n-2x-1}{n} \right]^{1/2},$$

$$0 \leq x \leq n-1$$

などが挙げられる(詳しくは, 中村・平井(2012))。

2項確率 $B(x; n, 0.5)$ で正規確率 $\Phi(z)$ を近似するのだから, x の方程式 $B(x; n, 0.5) = \Phi(z)$ の整数解を求めるのが自然である。しかし, この方程式の整数解は一般には存在しないし, 陽な形では表せない。この解についての議論は Molenaar(1973) で詳細に論じられている。条件 (b), (c) より, n が十分大きな場合には, 全ての $\hat{x}+i$ ($i = -1, 0, 1, 2$) は, 方程式 $B(x; n, 0.5) = \Phi(z)$ の近似解である。従って, $B(\hat{x}; n, 0.5)$, $B(\hat{x}+1; n, 0.5)$ はそれぞれ, $\Phi(z_0)$, $\Phi(z_1)$ の近似解である。不等式 $z_{-k+1} < \dots < z_{-1} < z_0 \leq z < z_1 < \dots < z_k$ より, z に近い z_0 , z_1 の値 $\Phi(z_0)$, $\Phi(z_1)$ を用いて $\Phi(z)$ の値を近似することを考える。 $\Phi(z_0)$, $\Phi(z_1)$ は, それぞれ, $B(\hat{x}; n, 0.5)$, $B(\hat{x}+1; n, 0.5)$ で近似できるから, $\Phi(z)$ の値を 2項確率 $B(\hat{x}; n, 0.5)$, $B(\hat{x}+1; n, 0.5)$ で近似する方法を提案する(図2.2参照)。

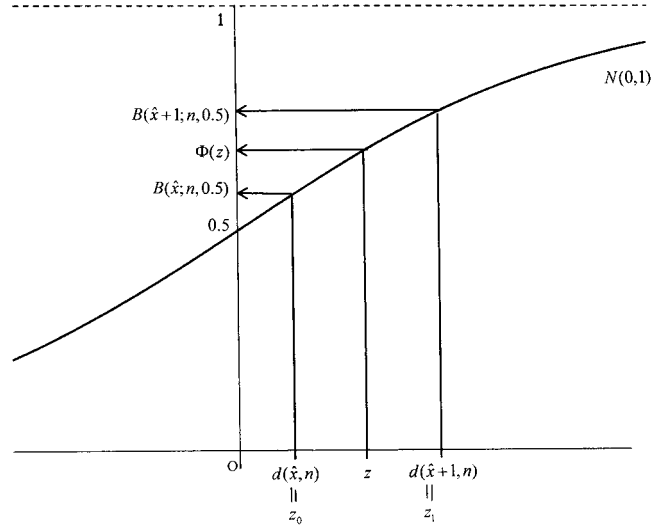


図 2.2. $\Phi(z), B(\hat{x}; n, 0.5), B(\hat{x}+1; n, 0.5)$ の位置関係

\tilde{x} , \tilde{z} を次式で定義する。

$$\tilde{x} = \begin{cases} \hat{x}, & 2z \leq z_1 + z_0, \\ \hat{x}+1, & 2z > z_1 + z_0; \end{cases} \quad \tilde{z} = d(\tilde{x}, n).$$

$t = \tilde{z}$ での $\Phi(t)$ の 0 次から 5 次までのテーラー展開

$$\Phi(t) = \Phi(\tilde{z}) + O(|t - \tilde{z}|),$$

$$\Phi(t) = \Phi(\tilde{z}) + \phi(\tilde{z})(t - \tilde{z}) + O(|t - \tilde{z}|^2),$$

$$\Phi(t) = \Phi(\tilde{z}) + \phi(\tilde{z})(t - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(t - \tilde{z}) \right] + O(|t - \tilde{z}|^3),$$

$$\Phi(t) = \Phi(\tilde{z}) + \phi(\tilde{z})(t - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(t - \tilde{z}) + \frac{1}{6}(\tilde{z}^2 - 1)(t - \tilde{z})^2 \right] + O(|t - \tilde{z}|^4),$$

$$\Phi(t) = \Phi(\tilde{z}) + \phi(\tilde{z})(t - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(t - \tilde{z}) + \frac{1}{6}(\tilde{z}^2 - 1)(t - \tilde{z})^2 + \frac{1}{24}(3\tilde{z} - \tilde{z}^3)(t - \tilde{z})^3 \right] + O(|t - \tilde{z}|^5),$$

$$\Phi(t) = \Phi(\tilde{z}) + \phi(\tilde{z})(t - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(t - \tilde{z}) + \frac{1}{6}(\tilde{z}^2 - 1)(t - \tilde{z})^2 + \frac{1}{24}(3\tilde{z} - \tilde{z}^3)(t - \tilde{z})^3 + \frac{1}{120}(3 - 6\tilde{z}^2 + \tilde{z}^4)(t - \tilde{z})^4 \right] + O(|t - \tilde{z}|^6)$$

を利用する。上の近似式はすべて未知の標準正規分布関数値 (i.e., 連続量) $\Phi(\tilde{z})$ と標準正規密度関数値 (i.e., 連続量) $\phi(\tilde{z})$ を含んでいる。本論では離散量のみを用いた正規確率計算アルゴリズムの開発を目的にしているため、これらを離散化する必要が生じる。第 2 節の条件 (c) より, n が十分大きいときは $\Phi(\tilde{z})$ を未知の量 $B(\tilde{x}; n, 0.5)$ で近似できる。 n が非常に大きい場合 (例えば, 数億) は, 通常の C 言語による数値処理プログラムでは $B(\tilde{x}; n, 0.5)$ を計算するのは困難である。平井・中村 (2006) は任意の $p \in (0, 1)$ と非常に広範囲の n お

よび任意の非負整数 x ($0 \leq x \leq n$) に対して, 2項確率 $B(x; n, p)$ の値の計算法を提案している。この計算法で $B(\tilde{x}; n, 0.5)$ は計算可能となる。この計算法によって算出された $B(\tilde{x}; n, 0.5)$ の値を $\hat{B}(\tilde{x}; n, 0.5)$ とかく。小規模の数値実験より, $\Phi(z)$ の近似値を求めた数値実験の結果, z の負の領域での相対誤差がよくない (例えば, 中村・平井(2012))。この欠点を補うために, 平井・中村(2006)の $B(x; n, p)$ の計算法を改善した計算法を新たに加える。修正計算法による $B(\tilde{x}; n, 0.5)$ の値を $\hat{C}_n(\tilde{x}) + \hat{B}(\tilde{x}; n, 0.5)$ とかく。この $B(x; n, p)$ の修正計算法は中村・平井(2012)に詳しく述べられている。一方, もう一つの連続量 $\phi(\tilde{z})$ の離散化について述べよう。ここでは, $t = \tilde{z}$ を含む3点での $\Phi(t)$ の近似値を利用して $\phi(\tilde{z})$ を近似する方法, いわゆる, 数値微分法を採用する。数値計算計算の分野では, 数値微分法として, Lagrange 補間法と差分近似法がよく知られている。ここでもこれらの方法による $\phi(\tilde{z})$ の近似 (数値微分という) を数種類採用し, これらに基づく正規確率の近似計算法を議論する。そのために若干の準備をする。

Case I: $2z \leq z_0 + z_1$ のとき, 3点 $(z_1, B(\hat{x}-1; n, 0.5))$, $(z_0, B(\hat{x}; n, 0.5))$, $(z_1, B(\hat{x}+1; n, 0.5))$ をとる。

Case II: $z_0 + z_1 < 2z$ のとき, 3点 $(z_0, B(\hat{x}; n, 0.5))$, $(z_1, B(\hat{x}+1; n, 0.5))$, $(z_2, B(\hat{x}+2; n, 0.5))$ をとる。

Case I の場合は $\tilde{z} = z_0$, Case II の場合は $\tilde{z} = z_1$ であることに注意しよう。表記を簡単にするため, Case I, Case II における3点を共通に表すのに記号 $(\tilde{z}_1, \tilde{B}_1)$, (\tilde{z}, \tilde{B}) , $(\tilde{z}_2, \tilde{B}_2)$ を用いる。これらの3点を利用して, $\Phi(z)$ の傾き $\phi(\tilde{z})$ の数値微分 $\hat{\phi}(\tilde{z})$ を求める。 $\hat{\phi}(\tilde{z})$ の数値微分として, いろいろな近似式がある (例えば, 中村・平井(2012))。簡単のため,

$$\hat{\phi}_-(\tilde{z}) = \frac{\tilde{B} - \tilde{B}_1}{\tilde{z} - \tilde{z}_1}, \quad \hat{\phi}_+(\tilde{z}) = \frac{\tilde{B}_2 - \tilde{B}}{\tilde{z}_2 - \tilde{z}}$$

とおく。ここでは次の近似式を採用する。

$$\hat{\phi}_1(\tilde{z}) = \begin{cases} \hat{\phi}_-(\tilde{z}), & z_0 + z_1 < 2z, \\ \hat{\phi}_+(\tilde{z}), & 2z \leq z_0 + z_1. \end{cases}$$

$$\hat{\phi}_2(\tilde{z}) = \begin{cases} \frac{\hat{\phi}_-(\tilde{z})}{1 - \frac{\tilde{z}}{2}(\tilde{z}_1 - \tilde{z})}, & z_0 + z_1 < 2z, \\ \frac{\hat{\phi}_+(\tilde{z})}{1 - \frac{\tilde{z}}{2}(\tilde{z}_2 - \tilde{z})}, & 2z \leq z_0 + z_1. \end{cases}$$

$$\hat{\phi}_3(\tilde{z}) = \begin{cases} \frac{\hat{\phi}_-(\tilde{z})}{1 - \frac{\tilde{z}}{2}(\tilde{z}_1 - \tilde{z}) + \frac{1}{6}(\tilde{z}^2 - 1)(\tilde{z}_1 - \tilde{z})^2}, & z_0 + z_1 < 2z, \\ \frac{\hat{\phi}_+(\tilde{z})}{1 - \frac{\tilde{z}}{2}(\tilde{z}_2 - \tilde{z}) + \frac{1}{6}(\tilde{z}^2 - 1)(\tilde{z}_2 - \tilde{z})^2}, & 2z \leq z_0 + z_1. \end{cases}$$

$$\hat{\phi}_4(\tilde{z}) = \frac{\tilde{z}_2 - \tilde{z}}{\tilde{z}_2 - \tilde{z}_1} \hat{\phi}_-(\tilde{z}) + \frac{\tilde{z} - \tilde{z}_1}{\tilde{z}_2 - \tilde{z}_1} \hat{\phi}_+(\tilde{z}).$$

上で述べたことの離散化により，次の近似式が得られる．

(I) 定数項 $B(\tilde{x}; n, 0.5)$ の計算に $\hat{B}(\tilde{x}; n, 0.5)$ を使用した場合：

$$\tilde{\Phi}_0(z) = \hat{B}(\tilde{x}; n, 0.5),$$

$$\tilde{\Phi}_{i1}(z) = \hat{B}(\tilde{x}; n, 0.5) + \hat{\phi}_i(\tilde{z})(z - \tilde{z}), \quad i = 1, 2, 3, 4,$$

$$\tilde{\Phi}_{i2}(z) = \hat{B}(\tilde{x}; n, 0.5) + \hat{\phi}_i(\tilde{z})(z - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(z - \tilde{z}) \right], \quad i = 1, 2, 3, 4,$$

$$\tilde{\Phi}_{i3}(z) = \hat{B}(\tilde{x}; n, 0.5) + \hat{\phi}_i(\tilde{z})(z - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(z - \tilde{z}) + \frac{1}{6}(\tilde{z}^2 - 1)(z - \tilde{z})^2 \right], \quad i = 1, 2, 3, 4,$$

$$\tilde{\Phi}_{i4}(z) = \hat{B}(\tilde{x}; n, 0.5) + \hat{\phi}_i(\tilde{z})(z - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(z - \tilde{z}) + \frac{1}{6}(\tilde{z}^2 - 1)(z - \tilde{z})^2 + \frac{1}{24}(3\tilde{z} - \tilde{z}^3)(z - \tilde{z})^3 \right], \quad i = 1, 2, 3, 4,$$

$$\tilde{\Phi}_{i5}(z) = \hat{B}(\tilde{x}; n, 0.5) + \hat{\phi}_i(\tilde{z})(z - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(z - \tilde{z}) + \frac{1}{6}(\tilde{z}^2 - 1)(z - \tilde{z})^2 + \frac{1}{24}(3\tilde{z} - \tilde{z}^3)(z - \tilde{z})^3 + \frac{1}{120}(3 - 6\tilde{z}^2 + \tilde{z}^4)(z - \tilde{z})^4 \right], \quad i = 1, 2, 3, 4.$$

(II) 定数項 $B(\tilde{x}; n, 0.5)$ の計算に $\hat{C}_n(\tilde{x}) + \hat{B}(\tilde{x}; n, 0.5)$ を使用した場合：

$$\hat{\Phi}_0(z) = \hat{C}_n(\tilde{x}) + \hat{B}(\tilde{x}; n, 0.5),$$

$$\hat{\Phi}_{i1}(z) = \hat{C}_n(\tilde{x}) + \hat{B}(\tilde{x}; n, 0.5) + \hat{\phi}_i(\tilde{z})(z - \tilde{z}), \quad i = 1, 2, 3, 4,$$

$$\hat{\Phi}_{i2}(z) = \hat{C}_n(\tilde{x}) + \hat{B}(\tilde{x}; n, 0.5) + \hat{\phi}_i(\tilde{z})(z - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(z - \tilde{z}) \right], \quad i = 1, 2, 3, 4,$$

$$\hat{\Phi}_{i3}(z) = \hat{C}_n(\tilde{x}) + \hat{B}(\tilde{x}; n, 0.5) + \hat{\phi}_i(\tilde{z})(z - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(z - \tilde{z}) + \frac{1}{6}(\tilde{z}^2 - 1)(z - \tilde{z})^2 \right],$$

$$i = 1, 2, 3, 4,$$

$$\hat{\Phi}_{i4}(z) = \hat{C}_n(\tilde{x}) + \hat{B}(\tilde{x}; n, 0.5) + \hat{\phi}_i(\tilde{z})(z - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(z - \tilde{z}) + \frac{1}{6}(\tilde{z}^2 - 1)(z - \tilde{z})^2 + \frac{1}{24}(3\tilde{z} - \tilde{z}^3)(z - \tilde{z})^3 \right], \quad i = 1, 2, 3, 4,$$

$$\hat{\Phi}_{i5}(z) = \hat{C}_n(\tilde{x}) + \hat{B}(\tilde{x}; n, 0.5) + \hat{\phi}_i(\tilde{z})(z - \tilde{z}) \left[1 - \frac{\tilde{z}}{2}(z - \tilde{z}) + \frac{1}{6}(\tilde{z}^2 - 1)(z - \tilde{z})^2 + \frac{1}{24}(3\tilde{z} - \tilde{z}^3)(z - \tilde{z})^3 + \frac{1}{120}(3 - 6\tilde{z}^2 + \tilde{z}^4)(z - \tilde{z})^4 \right], \quad i = 1, 2, 3, 4.$$

上の42種類の近似式と n の組み合わせによる無限個の近似計算法が得られたことになる。標準正規関数の計算法としてこれらの近似式のどれを選べばよいのかという問題が生じる。これについては次節で議論する。

3. 近似法の集合を基にした正規確率計算法の構成手続き

第2節でいろいろな近似法を導入した。そこで構成される近似法は分割関数, 5次までの多項式, 定数項の2項近似法, 自然数のパラメータの4種類の属性を含んでいる。従って, 無限個の近似法があり, これらの近似法群に精度の良さを測る基準を設け, その基準の下で“良い近似法を構成する”必要がある。定数項の近似計算法について述べる。平井・中村(2006)が提案した2項近似計算法よりも中村・平井(2012)が提案した修正2項近似計算法の方が精度が良いことが, 小規模の数値実験により, 示されている。この理由により, 定数項の2項近似計算法として修正2項近似計算法を採用する。自然数のパラメータ n についても, 小規模の数値実験の結果より, $n=10^9$ の場合はそれ以下のパラメータ n を持つ近似法より近似精度が悪いので, 10億 (10^9) 以下のパラメータ n を持つ近似法に制限した。 n が大きくなると近似法の精度が悪くなる理由として, n が大きくなると計算量が多くなり, コンピュータ内部計算において誤差が積み重なってくるものと推察される。一方, n が小さいと近似誤差は悪い。ここでは近似誤差が 10^{-3} 以下であるような n をもつ近似計算法に制限する。ド・モアブル=ラプラスの極限定理における近似誤差は $1/\sqrt{n}$ のオーダー(例えば, Johnson, Kotz and Balakrishnan(1994))だから, $\sqrt{n} \geq 10^3$ を満たすような大きさの n に制限する。すなわち, 100万 (10^6) 以上のパラメータ n を持つ近似法に制限する。以上の理由により, 良い近似法を構成するための土台となる近似法の全体として,

$$\mathcal{A} = \{\hat{\Phi}_k(t; n, d, \hat{\phi}_i); n \in \mathcal{N} \text{ and } 10^6 \leq n \leq 10^9, d = d_{CN}, d_{CF}, d_{PP}, i = 1, 2, 3, 4, k = 0, 1, 2, 3, 4, 5\}$$

を採用する。ここに, \mathcal{N} は自然数の全体である。 $\hat{\Phi}_k(t; n, d, \hat{\phi}_i)$ の個々の具体的な形は第2節に示されているが, 定義より $\hat{\Phi}_0(t; n, d, \hat{\phi}_1) = \hat{\Phi}_0(t; n, d, \hat{\phi}_2) = \hat{\Phi}_0(t; n, d, \hat{\phi}_3) = \hat{\Phi}_0(t; n, d, \hat{\phi}_4)$ である。以下, 良い近似法を構成するための土台となる近似法の全体(例えば, \mathcal{A}) を“土台集合”と呼ぶことにする。

3.1 比較のための基準の設定

集合 \mathcal{A} に属す近似法の精度を比較するために, 近似法 $f \in \mathcal{A}$ の点 $t \in \mathcal{R}$ における絶対相対誤差 $|f(t) - \Phi(t)| / \Phi(t)$ を利用する。 \mathcal{R} の部分区間 I における近似法 $f \in \mathcal{A}$ の評価 $ARE(f; I)$ を $\max_{t \in I} |f(t) - \Phi(t)| / \Phi(t)$ で定義する。通常のC言語とパソコンで得られる精度は 10^{-15} ないしは 10^{-16} 程度である。従って, \mathcal{R} の全域において近似法 $f \in \mathcal{A}$ の評価 $ARE(f; \mathcal{R})$ を計算することは, 精度の観点から, 無駄である。このため, 近似法 $f \in \mathcal{A}$ の評価の計算する範囲として, パソコンで計算できる精度の範囲を考えれば十分である。よく知られた不等式 (Abramowitz and Stegun, 1972)

$$\frac{\exp(-t^2)}{t + \sqrt{t^2 + 2}} < \int_t^\infty \exp(-t^2) dt < \frac{\exp(-t^2)}{t + \sqrt{t^2 + 4/\pi}}, \quad t \geq 0$$

より,

$$\frac{1}{\sqrt{\pi}} \frac{\exp(-t^2)}{t + \sqrt{t^2 + 2}} < 1 - \Phi(\sqrt{2}t) < \frac{1}{\sqrt{\pi}} \frac{\exp(-t^2)}{t + \sqrt{t^2 + 4/\pi}}$$

が得られる。上の不等式に $t = 8/\sqrt{2}$ を代入すると, $0.62196 \cdots \times 10^{-15} < 1 - \Phi(8) < 0.62537 \cdots \times 10^{-15}$ となる。Cパソコンの持つ精度 (10^{-15} ないしは 10^{-16} 程度) より, 近似法 $f \in \mathcal{A}$ の評価を計算する範囲として区間 $[-8, 8]$ を採用する。

条件

$$ARE(g; [-8, 8]) = \min\{ARE(f; [-8, 8]); f \in \mathcal{A}\}$$

を満たす $g \in \mathcal{A}$ は区間 $[-8, 8]$ が広すぎて, 絶対相対誤差の基準で良いと判断できる近似法が特定できない。このため, 区間 $[-8, 8]$ の部分区間において, 何らかの選択基準を設定する必要がある。区間 $[-8, 8]$ の部分区間 I において, $g \in \mathcal{A}$ より $f \in \mathcal{A}$ の方が精度が良いとは $ARE(f; I) \leq ARE(g; I)$ であるときにいう。部分区間 I において, $\hat{f} \in \mathcal{A}$ が最良の近似法と

$$ARE(\hat{f}; I) = \min\{ARE(f; I); f \in \mathcal{A}\}$$

のときにいう。集合 \mathcal{A} に属す近似法は解析的でないので, どの近似法の精度が良いのかを理論的に示すのは

困難である。このため、数値実験を通じてそれらの中から最良の近似法を構成する。具体的な数値計算の設定や近似法の良さ具合の基準を以下に述べる。

区間 $[-8,8]$ を部分区間 $I_1 = [-8,-6)$, $I_2 = [-6,-4)$, \dots , $I_7 = [4,6)$, $I_8 = [6,8]$ に分割する。 $[-8,8]$ の任意の部分区間 I において適当な基準を設ける。その基準を各部分区間 I_m ($1 \leq m \leq 8$) に対して適用し、良い近似法を選択する方法を導入する。 $ARE(f; I_m)$ ($1 \leq m \leq 8$) は計算できないので、近似値を利用する。そのために、区間 $[-8,8]$ を幅 10^{-2} の微小区間に等分割し、それぞれの微小区間の中点をとる。即ち、分割の左側から j 番目の微小区間の中点 t_j の値は、

$$t_j = \frac{2j-1-1600}{200}, \quad j=1, 2, \dots, 1600$$

となる。これらの1600個の点での各近似法の絶対相対誤差をモグラ算術で計算する。真値 $\Phi(t_j)$ の値は Maple の50桁精度計算で算出した数を採用した。計算は NEC 社製の LaVieLL750(Intel(R) Core(TM) i5 CPU 2.53GHz, 実装メモリ 4GB, 64ビット Windows 7) で実行し、コンパイラは Microsoft Visual C++ 2010を用いた。区間 $[-8,8]$ の部分区間 I における近似法 $f \in \mathcal{A}$ の評価基準 $ARE(f; I)$ の近似値として、

$$\widehat{ARE}(f; I) = \max\left\{\frac{|f(t_j) - \Phi(t_j)|}{\Phi(t_j)}; t_j \in I, 1 \leq j \leq 1600\right\}$$

を採用する。 $I = [-8,8]$ の場合は単に $\widehat{ARE}(f)$ とかくことにする。以下、混乱の恐れがない限り、記号 $\widehat{ARE}(f; I)$ の $\widehat{}$ は省略することにする。

3.2 最良近似法の選出アルゴリズム

選出アルゴリズムを述べる前に、若干の記号を導入する。 $\#\mathcal{A}$ は非常に大きいので、 n として、10万から1000万まで10万ずつ増やした値、2000万から3億までは1000万ずつ増やした値、4億から10億までは1億ずつ増やした値、さらに15億と20億を採用する。これらの n の全体を

$$\mathcal{N}_0 = \{10^5 k; k=1, \dots, 99\} \cup \{10^7 k; k=1, \dots, 29\} \cup \{10^8 k; k=3, \dots, 9\} \cup \{10^9 + 5 \times 10^8 k; k=0, 1, 2\}$$

とおく。良い近似法を構成するための最初の土台集合として

$$\mathcal{A}_0 = \{\hat{\Phi}_k(t; n, d, \hat{\phi}_i); n \in \mathcal{N}_0, d = d_{CN}, d_{CF}, d_{PP}, i=1, 2, 3, 4, k=0, 1, 2, 3, 4, 5\}$$

を採用する。区間 $[-8,8]$ の任意の部分区間 I に対して、

$$a(n, d, i, k; I) = \max_{1 \leq j \leq 1000} ARE(\hat{\Phi}_k(t_j; n, d, \hat{\phi}_i); I), \quad n \in \mathcal{N},$$

$$\bar{a}(n, d, i, k; I) = \frac{a(n-1, d, i, k; I) + a(n, d, i, k; I)}{2},$$

$$\mathcal{E}_0(I) = \{\bar{a}(n, d, i, k; I); n \in \mathcal{N}_0, d = d_{CN}, d_{CF}, d_{PP}, i=1, 2, 3, 4, k=0, 1, 2, 3, 4, 5\}$$

とおく。データセット $\mathcal{E}_0(I)$ は土台集合 \mathcal{A}_0 を部分区間上の近似法の集合としてみたときの別形である。

ここで、 $\hat{\Phi}_k(t_j; n-1, d, \hat{\phi}_i)$ と $\hat{\Phi}_k(t_j; n, d, \hat{\phi}_i)$ を組にして扱う理由を述べよう。偶数の場合は数値計算に用いる点在各小区間の端点となる場合があり得るが、奇数の場合にはそれは起きない。そのために偶数と奇数では異なった精度になるので、偶数・奇数が精度の安定性に及ぼす影響をみる必要がある。後述のように、パラメータが偶数である近似法とパラメータが奇数である近似法を同時に考慮した近似法の構成が必然的に生じるのである。

$\bar{a}(n, d, i, k; I)$ が $\mathcal{E}_0(I)$ の最小値のとき、

$$\hat{\Phi} = \frac{1}{2}(\hat{\Phi}_k(t_j; n, d, \hat{\phi}_i) + \hat{\Phi}_k(t_j; n-1, d, \hat{\phi}_i))$$

を \mathcal{A}_0 を基にして構成した近似法という。この近似法は、

$$ARE(\hat{\Phi}) \leq \frac{1}{2}(ARE(\hat{\Phi}_k(t_j; n-1, d, \hat{\phi}_i)) + ARE(\hat{\Phi}_k(t_j; n, d, \hat{\phi}_i)))$$

と評価されるので、同程度の絶対相対誤差が期待される。

本論で提案する“良い近似法”の構成手順は、データセット $\mathcal{E}_0(I)$ を出発点として、段階的に土台集合を狭めながら良い近似法を構成していく方法である。最初に、第1段階の手順を述べよう。

ステップ1-0: $\mathcal{E} = \mathcal{E}_0(I)$ とする。

ステップ1-1: $\min\{v; v \in \mathcal{E}\}$ を $r \times 10^{-m}$ と表す。ただし、 m は非負整数、実数 r は $1 \leq r < 10$ を満たすものとする。 $\tilde{v} = (r+5.5) \times 10^{-m}$ とする。

ステップ1-2: 固定した d に対し、 $\bar{a}(n, d, i, k; I) \in \{v \in \mathcal{E}_0(I); v \leq \tilde{v}\}$ となる近似法の全体を $\mathcal{A}(d)$ とおく。 $\#(\mathcal{A}(d_{CN})), \#(\mathcal{A}(d_{CF})), \#(\mathcal{A}(d_{PP}))$ の中で最も大きいものに対応する分割関数を d^* とする。

ステップ1-3: $\hat{\Phi}_k(t; n, d^*, \hat{\phi}_i) \in \mathcal{A}(d^*)$ を満たす $n \in \mathcal{N}_0$ の最小値を n' とする。 $n < n'$ を満たす $n \in \mathcal{N}_0$ の最大値を n'' とする。

ステップ1-4:

- (i) $n' - n'' \leq 10^6$ ならば、 $\tilde{\mathcal{A}}(I) = \mathcal{A}(d^*)$ とおいて手続きを終了する。
- (ii) $n' - n'' \leq 10^6$ ならば、

$$\mathcal{N}' = \{n'' + (m-1) \times \frac{n' - n''}{10}; m = 1, 2, \dots, 10, 11\},$$

$$\mathcal{E}' = \{\bar{a}(n, d^*, i, k; I); n \in \mathcal{N}', i = 1, 2, 3, 4, k = 0, 1, 2, 3, 4, 5\}$$

とする。

ステップ1-5: $\mathcal{E} = \mathcal{E}'$ とおいて、ステップ1-0に戻る。

以上で第1段階の手続きが終了する。第1段階の手続きで得られた $\hat{\mathcal{A}}(I)$ に対して、次の第2段階の手続きを行う。

ステップ2-1: $\hat{\mathcal{A}}(I)$ に属す近似法の中で、パラメータが一番小さいものを選び、その全体を $\mathcal{A}'(I)$ とおく。

ステップ2-2: $\mathcal{A}'(I)$ に属す近似法の中で、多項式の次数が一番小さいものを選ぶ。その全体を $\tilde{\mathcal{A}}(I)$ とおく。

以上で、第2段階の手続きが終了する。第2段階の手続きで得られた $\tilde{\mathcal{A}}(I)$ に対して、次の第3段階の手続きを行う。

ステップ3-1: $\tilde{\mathcal{A}}(I)$ から任意に $\hat{\Phi}_k(t; n-1, d^*, \hat{\phi}_i)$ と $\hat{\Phi}_k(t; n, d^*, \hat{\phi}_i)$ を選び、

$$\hat{\Phi}(t; I) = \frac{1}{2}(\hat{\Phi}_k(t; n-1, d^*, \hat{\phi}_i) + \hat{\Phi}_k(t; n, d^*, \hat{\phi}_i))$$

とする。

以上で、第3段階の手続きが終了する。上記の3段階の手続きを経て、部分区間 I での標準正規分布関数 $\Phi(t)$ の近似計算法 $\hat{\Phi}(t; I)$ が得られる。この節で提案した手続きの応用については次節で議論する。

3.3 数値実験による正規確率近似法の構成

前節では、正規確率を近似するアルゴリズムの構成法を述べた。ここでは、数値実験により、正規確率の近似法を構成しよう。対象区間 $[-8, 8]$ を部分区間 $I_1 = [-8, -6)$, $I_2 = [-6, -4)$, \dots , $I_7 = [4, 6)$, $I_8 = [6, 8]$ に分割して、それぞれの部分区間で正規確率を近似するアルゴリズムを構成する。

例として、部分区間 $I_1 = [-8, -6)$ について詳細に各ステップの結果を示す。

[第1段階]

ステップ1-0: $\mathcal{E} = \mathcal{E}_0(I_1)$ とおく。

ステップ1-1: $\min\{v; v \in \mathcal{E}\} = 4.5376 \times 10^{-13}$, $\tilde{v} = 10.0376 \times 10^{-13}$

ステップ1-2 :

表3.1に示すのは, $\bar{a}(n, d, i, k; I) \in \{v \in \mathcal{E}_0(I); v \leq \tilde{v}\}$ となるすべての近似法の $\bar{a}(n, d, i, k; I)$ の値である。

表 3. 1.

$\bar{a}(2\text{億}, d_{CF}, 3, 3; I)$	4. 5376E-13
$\bar{a}(2\text{億}, d_{CF}, 3, 4; I)$	4. 5398E-13
$\bar{a}(2\text{億}, d_{CF}, 3, 5; I)$	4. 5398E-13
$\bar{a}(4\text{億}, d_{CF}, 3, 3; I)$	5. 1270E-13
$\bar{a}(4\text{億}, d_{CF}, 3, 4; I)$	5. 1270E-13
$\bar{a}(4\text{億}, d_{CF}, 3, 5; I)$	5. 1318E-13
$\bar{a}(1\text{億}, d_{CF}, 3, 3; I)$	6. 7081E-13
$\bar{a}(1\text{億}, d_{CF}, 3, 4; I)$	6. 7342E-13
$\bar{a}(1\text{億}, d_{CF}, 3, 5; I)$	6. 7342E-13

これより, $\#(\mathcal{E}(d_{CN})) = 0$, $\#(\mathcal{E}(d_{CF})) = 9$, $\#(\mathcal{E}(d_{PP})) = 0$ 。よって, 分割関数 d_{CF} が選ばれる。

ステップ1-3 : $n' = 1\text{億}$, $n'' = 9000\text{万}$ となる。

ステップ1-4 : $n' - n'' > 10^6$ なので,

$$\mathcal{N}' = \{9000\text{万}, 9100\text{万}, \dots, 1\text{億}\},$$

$$\mathcal{E}' = \{\bar{a}(n, d_{CF}, i, k; I); n \in \mathcal{N}', i = 1, 2, 3, 4, k = 0, 1, 2, 3, 4, 5\}$$

ステップ1-5 : $\mathcal{E} = \mathcal{E}'$ とにおいて, ステップ1-0に戻る。

ステップ1-0 : $\mathcal{E} = \mathcal{E}_0(I_1)$ とおく。

ステップ1-1 : $\min\{v; v \in \mathcal{E}\} = 6.7081 \times 10^{-13}$, $\tilde{v} = 12.2081 \times 10^{-13}$

ステップ1-2 : 表3.2に示すのは, $\bar{a}(n, d, i, k; I) \in \{v \in \mathcal{E}_0(I); v \leq \tilde{v}\}$ となるすべての近似法の $\bar{a}(n, d, i, k; I)$ の値である。

表 3. 2.

$\bar{a}(1\text{億}, d_{CF}, 3, 3; I)$	6. 7081E-13	$\bar{a}(9800\text{万}, d_{CF}, 3, 3; I)$	1. 0615E-12
$\bar{a}(1\text{億}, d_{CF}, 3, 4; I)$	6. 7342E-13	$\bar{a}(9800\text{万}, d_{CF}, 3, 4; I)$	1. 0615E-12
$\bar{a}(1\text{億}, d_{CF}, 3, 5; I)$	6. 7342E-13	$\bar{a}(9800\text{万}, d_{CF}, 3, 5; I)$	1. 0615E-12
$\bar{a}(9900\text{万}, d_{CF}, 3, 3; I)$	9. 9194E-13	$\bar{a}(9700\text{万}, d_{CF}, 3, 3; I)$	1. 1598E-12
$\bar{a}(9900\text{万}, d_{CF}, 3, 4; I)$	9. 9194E-13	$\bar{a}(9700\text{万}, d_{CF}, 3, 4; I)$	1. 1598E-12
$\bar{a}(9900\text{万}, d_{CF}, 3, 5; I)$	9. 9194E-13	$\bar{a}(9700\text{万}, d_{CF}, 3, 5; I)$	1. 1598E-12
$\bar{a}(9300\text{万}, d_{CF}, 3, 3; I)$	1. 0186E-12	$\bar{a}(9000\text{万}, d_{CF}, 3, 3; I)$	1. 2152E-12
$\bar{a}(9300\text{万}, d_{CF}, 3, 4; I)$	1. 0186E-12	$\bar{a}(9000\text{万}, d_{CF}, 3, 4; I)$	1. 2152E-12
$\bar{a}(9300\text{万}, d_{CF}, 3, 5; I)$	1. 0186E-12	$\bar{a}(9000\text{万}, d_{CF}, 3, 5; I)$	1. 2152E-12

ステップ1-3 : $n' = 9000\text{万}$, $n'' = 8000\text{万}$ となる。

ステップ1-4 : $n' - n'' > 10^6$ なので,

$$\mathcal{N}' = \{8000\text{万}, 8100\text{万}, \dots, 9000\text{万}\},$$

$$\mathcal{E}' = \{\bar{a}(n, d_{CF}, i, k; I); n \in \mathcal{N}', i = 1, 2, 3, 4, k = 0, 1, 2, 3, 4, 5\}$$

ステップ1-5 : $\mathcal{E} = \mathcal{E}'$ とにおいて, ステップ1-0に戻る。

ステップ1-0 : $\mathcal{E} = \mathcal{E}_0(I_1)$ とおく。

ステップ1-1 : $\min\{v; v \in \mathcal{C}\} = 8.2580 \times 10^{-13}$, $\tilde{v} = 13.7579 \times 10^{-13}$

ステップ1-2 : 表3.3に示すのは, $\bar{a}(n, d, i, k; I) \in \{v \in \mathcal{C}_0(I); v \leq \tilde{v}\}$ となるすべての近似法の $\bar{a}(n, d, i, k; I)$ の値である。

表 3.3.

$\bar{a}(8800\text{万}, d_{CF}, 3, 3; I)$	8.2579E-13	$\bar{a}(8700\text{万}, d_{CF}, 3, 3; I)$	9.6206E-13
$\bar{a}(8800\text{万}, d_{CF}, 3, 4; I)$	8.2579E-13	$\bar{a}(8700\text{万}, d_{CF}, 3, 4; I)$	9.6206E-13
$\bar{a}(8800\text{万}, d_{CF}, 3, 5; I)$	8.2579E-13	$\bar{a}(8700\text{万}, d_{CF}, 3, 5; I)$	9.6206E-13
$\bar{a}(8900\text{万}, d_{CF}, 3, 3; I)$	8.8140E-13	$\bar{a}(8600\text{万}, d_{CF}, 3, 3; I)$	1.1474E-13
$\bar{a}(8900\text{万}, d_{CF}, 3, 4; I)$	8.8140E-13	$\bar{a}(8600\text{万}, d_{CF}, 3, 4; I)$	1.1474E-12
$\bar{a}(8900\text{万}, d_{CF}, 3, 5; I)$	8.8148E-13	$\bar{a}(8600\text{万}, d_{CF}, 3, 5; I)$	1.1474E-12
$\bar{a}(8500\text{万}, d_{CF}, 3, 3; I)$	9.4960E-13	$\bar{a}(9000\text{万}, d_{CF}, 3, 3; I)$	1.2152E-12
$\bar{a}(8500\text{万}, d_{CF}, 3, 4; I)$	9.4960E-13	$\bar{a}(9000\text{万}, d_{CF}, 3, 4; I)$	1.2152E-12
$\bar{a}(8500\text{万}, d_{CF}, 3, 5; I)$	9.4960E-13	$\bar{a}(9000\text{万}, d_{CF}, 3, 5; I)$	1.2152E-12

ステップ1-3 : $n' = 8500\text{万}$, $n'' = 8400\text{万}$ となる。

ステップ1-4 : $n' - n'' \leq 10^6$ なので, $\tilde{\mathcal{A}}(I) = \mathcal{A}(d_{CF})$

[第2段階]

ステップ2-1 : $\mathcal{A}'(I)$ に属す近似法の $\bar{a}(n, d, i, k; I)$ は,

$$\bar{a}(8500\text{万}, d_{CF}, 3, 3; I), \bar{a}(8500\text{万}, d_{CF}, 3, 4; I), \bar{a}(8500\text{万}, d_{CF}, 3, 5; I)$$

となる。

ステップ2-2 : $\hat{\mathcal{A}}(I)$ に属す近似法の $\bar{a}(n, d, i, k; I)$ は, $\bar{a}(8500\text{万}, d_{CF}, 3, 3; I)$ となる。

[第3段階]

ステップ3-1 : $\hat{\Phi}(t; I) = \frac{1}{2}(\hat{\Phi}_3(t_j; 8500\text{万} - 1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t_j; 8500\text{万}, d_{CF}, \hat{\phi}_3))$

が得られる。

このとき, $ARE(\frac{1}{2}(\hat{\Phi}_3(t_j; n-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t_j; n, d_{CF}, \hat{\phi}_3))) = 8.9906\text{E}-13$ である。

各部分区間 I_2, \dots, I_8 についても, 同様に近似法の構成をおこなう。最終的に, I_1, I_2, \dots, I_8 の各部分区間で構成される近似法は, 表3.4の通りである。

表 3.4. 各部分区間 I_1, I_2, \dots, I_8 で構成される近似法

部分区間	提案される近似法	左から順に $ARE(\hat{\Phi}_3(t; n-1, d_{CF}, \hat{\phi}_3)),$ $ARE(\hat{\Phi}_3(t; n, d_{CF}, \hat{\phi}_3)),$ $ARE(\frac{1}{2}(\hat{\Phi}_3(t; n-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; n, d_{CF}, \hat{\phi}_3)))$		
$[-8, -6)$	$\frac{1}{2}(\hat{\Phi}_3(t; 8500万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 8500万, d_{CF}, \hat{\phi}_3))$	1. 2232E-12	6. 7602E-13	8. 9906E-13
$[-6, -4)$	$\frac{1}{2}(\hat{\Phi}_3(t; 8500万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 8500万, d_{CF}, \hat{\phi}_3))$	3. 2359E-13	3. 5596E-13	9. 6200E-14
$[-4, -2)$	$\frac{1}{2}(\hat{\Phi}_3(t; 900万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 900万, d_{CF}, \hat{\phi}_3))$	1. 2249E-13	3. 6608E-13	7. 9549E-13
$[-2, 0)$	$\frac{1}{2}(\hat{\Phi}_3(t; 400万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 400万, d_{CF}, \hat{\phi}_3))$	8. 2961E-14	6. 6940E-14	4. 9868E-14
$[0, 2)$	$\frac{1}{2}(\hat{\Phi}_3(t; 400万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 400万, d_{CF}, \hat{\phi}_3))$	2. 4041E-15	2. 1153E-14	1. 1146E-14
$[2, 4)$	$\frac{1}{2}(\hat{\Phi}_3(t; 300万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 300万, d_{CF}, \hat{\phi}_3))$	6. 4554E-15	5. 4931E-15	4. 2035E-15
$[4, 6)$	$\frac{1}{2}(\hat{\Phi}_3(t; 300万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 300万, d_{CF}, \hat{\phi}_3))$	4. 4410E-16	4. 4410E-16	3. 3308E-16
$[6, 8]$	$\frac{1}{2}(\hat{\Phi}_3(t; 100万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 100万, d_{CF}, \hat{\phi}_3))$	1. 1102E-16	1. 1102E-16	1. 1102E-16

このうち、部分区間 $[6, 8]$ では、近似法の構成の過程で、 $n < 10^6$ のパラメータ n をもつ近似法も最後まで残ったが、第 3 節に述べた理由により、 $n = 100$ 万が選ばれた。

3.4 提案された近似法の検証

提案された近似法が別の $t \in [-8, 8]$ についても同じ精度を示すことを検証する。前節同様に部分区間 $I_1 = [-8, -6)$, $I_2 = [-6, -4)$, \dots , $I_7 = [4, 6)$, $I_8 = [6, 8]$ に分割する。区間 $[-8, 8]$ を幅 10^{-3} の微小区間に等分割し、それぞれの微小区間内に乱数を用いて 1 点をとる。このようにしてとらえた 16000 個の点に対して、前節で構成された近似法を実行する。各部分区間 I_1, \dots, I_8 での $ARE(\frac{1}{2}(\hat{\Phi}_3(t_j; n-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t_j; n, d_{CF}, \hat{\phi}_3)))$ は表 3.5 に示す通りである。前節の表 3.4 で示された $ARE(\frac{1}{2}(\hat{\Phi}_3(t_j; n-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t_j; n, d_{CF}, \hat{\phi}_3)))$ と同程度の値であることがわかる。

表 3.5. 各部分区間 I_1, I_2, \dots, I_8 で提案される近似法

部分区間	提案される近似法	$ARE\left(\frac{1}{2}(\hat{\Phi}_3(t; n-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; n, d_{CF}, \hat{\phi}_3))\right)$
[-8, -6)	$\frac{1}{2}(\hat{\Phi}_3(t; 8500万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 8500万, d_{CF}, \hat{\phi}_3))$	9.0813E-13
[-6, -4)	$\frac{1}{2}(\hat{\Phi}_3(t; 8500万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 8500万, d_{CF}, \hat{\phi}_3))$	9.7293E-14
[-4, -2)	$\frac{1}{2}(\hat{\Phi}_3(t; 900万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 900万, d_{CF}, \hat{\phi}_3))$	7.8797E-13
[-2, 0)	$\frac{1}{2}(\hat{\Phi}_3(t; 400万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 400万, d_{CF}, \hat{\phi}_3))$	4.8015E-14
[0, 2)	$\frac{1}{2}(\hat{\Phi}_3(t; 400万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 400万, d_{CF}, \hat{\phi}_3))$	1.1105E-14
[2, 4)	$\frac{1}{2}(\hat{\Phi}_3(t; 300万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 300万, d_{CF}, \hat{\phi}_3))$	4.4286E-15
[4, 6)	$\frac{1}{2}(\hat{\Phi}_3(t; 300万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 300万, d_{CF}, \hat{\phi}_3))$	4.4410E-16
[6, 8]	$\frac{1}{2}(\hat{\Phi}_3(t; 100万-1, d_{CF}, \hat{\phi}_3) + \hat{\Phi}_3(t; 100万, d_{CF}, \hat{\phi}_3))$	1.1102E-16

4. 従来近似法との比較

正規分布関数を近似する従来の手法をいくつか取り上げて今回の方法との比較を試みよう。以下に示すのは、展開式による方法と近似式による方法での精度である（例えば、山内（1992））。絶対誤差 ε は（計算値）－（真値）、相対誤差は（（計算値）－（真値））／（真値）である。ここでは（真値）は正規分布関数の値であるから（真値） ≤ 1 である。よって、（絶対誤差 ε の絶対値） \leq （相対誤差の絶対値）となる。以下に述べる従来の手法との比較では、（絶対誤差 ε の絶対値） \leq （相対誤差の絶対値）の関係をふまえて議論をする。

まず、本論で提案する方法での最大絶対誤差と最大絶対相対誤差は以下の通りである。

絶対誤差 $|\varepsilon|$: 区間 [-8, 8] で、 $|\varepsilon| < 1.3989 \times 10^{-14}$

絶対相対誤差 : 区間 [-8, 8] で、最大絶対相対誤差が 9.0813×10^{-13} 。

以下、(1) から (6) で従来の方法との比較を述べる。

(1) マクローリン展開による方法

$$\Phi(t) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \left[t - \frac{t^2}{2 \cdot 1! \cdot 3} + \frac{t^5}{2^2 \cdot 2! \cdot 5} - \dots + (-1)^k \frac{t^{2k+1}}{2^k \cdot k! \cdot (2k+1)} + \dots \right]$$

絶対誤差 : $k = 4$ のとき、 $t = 0.1$ で 9.2×10^{-17} 、 $t = 1.0$ で 8.8×10^{-6}

絶対相対誤差 : 記載なし

提案する方法との比較 : t が小さいときはこの方法の方が精度が高いが、 t が大きくなると提案する方法の方が著しく精度が高い。

(2) 連分数展開によるもの (Shenton, 1954)

$$\Phi(t) = \frac{1}{2} + \phi(t) \left[\frac{t}{1} - \frac{t^2}{3} + \frac{2t^2}{5} - \dots + (-1)^k \frac{kt^2}{2k+1} + \dots \right]$$

絶対誤差： $k = 4$ のとき， $t = 0.1$ で 5.1×10^{-17} ， $t = 1.0$ で 4.7×10^{-6}

絶対相対誤差：記載なし

提案する方法との比較： t が小さいときはこの方法の方が精度が高いが， t が大きくなると提案する方法の方が著しく精度が高い。

(3) 漸近展開式による方法

$$\Phi(t) \doteq \frac{1}{t} \phi(t) \left[1 - \frac{1}{t^2} + \frac{1 \cdot 3}{t^4} - \frac{1 \cdot 3 \cdot 5}{t^6} + \dots + (-1)^k \frac{1 \cdot 3 \cdot \dots \cdot (2k-1)}{t^{2k}} + \dots \right], \quad t > 0$$

絶対誤差： $k = 5$ のとき， $t = 2.0$ で 1.8×10^{-2} ， $t = 5.0$ で 8.5×10^{-12}

絶対相対誤差：記載なし

提案する方法との比較： t が大きいときでも提案する方法の方が少し精度が高い。さらに t が小さくなると，この方法では精度が大変低くなる。

(4) 近似式による方法 (Hastings, 1955)

$$\Phi(t) \doteq 1 - \frac{1}{2} (1 + d_1 t + d_2 t^2 + d_3 t^3 + d_4 t^4 + d_5 t^5 + d_6 t^6)^{-16}, \quad t \geq 0$$

$$d_1 = 0.0498673470, d_2 = 0.0211410061, d_3 = 0.0032776263,$$

$$d_4 = 0.0000380036, d_5 = 0.0000488906, d_6 = 0.0000053830$$

絶対誤差 $|\varepsilon|$ ： $t \geq 0$ の範囲で， $|\varepsilon| < 1.3 \times 10^{-7}$

絶対相対誤差：記載なし

提案する方法との比較： $t \geq 0$ の範囲のみで比較しても，提案する方法の方がかなり精度が高い。

(5) 近似式による方法 (山内, 1964)

$$\Phi(t) \doteq \frac{1}{2} + \frac{1}{2} \left[1 - e^{-\frac{2t^2}{\pi}} \left\{ 1 + t^4 \left(0.0055 + \frac{0.0551}{t^2 + 14.4} \right) \right\} \right]^{\frac{1}{2}}, \quad t \geq 0$$

絶対誤差 $|\varepsilon|$ ： $t \geq 0$ の範囲で， $|\varepsilon| < 2.04 \times 10^{-5}$

絶対相対誤差：記載なし

提案する方法との比較： $t \geq 0$ の範囲のみで比較しても，提案する方法の方がかなり精度が高い。

(6) 近似式による方法 (Williams- 山内, 1964)

$$\Phi(t) \doteq \frac{1}{2} + \frac{1}{2} \left[1 - e^{-\frac{2t^2}{\pi}} \left\{ 1 + \frac{2(\pi-3)}{3\pi^2} t^4 \right\} \right]^{\frac{1}{2}}, \quad t \geq 0$$

絶対誤差 $|\varepsilon|$ ： $t \geq 0$ の範囲で， $|\varepsilon| < 3.8 \times 10^{-4}$

絶対相対誤差：記載なし

提案する方法との比較： $t \geq 0$ の範囲のみで比較しても，提案する方法の方がかなり精度が高い。

以上，(1) から (6) までに従来の方法との比較をおこなったが，従来の方法は t の一部の範囲では優れた精度をもつ場合もあるが， $[-8, 8]$ のような広範囲の t について一様に良い精度をもつ手法ではないことがわかる。その意味で，今回提案された近似法は，絶対相対誤差基準で $[-8, 8]$ という広範囲の t について一様に良い精度をもつという点で優れていることがわかる。

5. おわりに

提案された近似法は幅広い区間で最大絶対相対誤差 9.0813×10^{-13} を持っていることが検証された。計算時間についても数値実験をおこなった結果，部分区間 $[-8, -6)$ では 1 点当たりの平均計算時間が 2.935 秒，部分区間 $[-2, 0)$ と $[0, 2)$ では 0.135 秒，部分区間 $[2, 4)$ では 0.100 秒等となった。従って，提案された近似法は計算

精度・計算時間の面からも十分実用的であるといえる。更に、何時でも誰でも利用できる様に、この近似法は Web 上から実行形式のファイルとしてダウンロード可能となっている（サイトアドレスは、http://www.cc.okayama-u.ac.jp/~qdca1385/stat2012/Binom_webStart.htm である）。

参考文献

- Abramowitz, M. and Stegun, I.A. (ed.) (1972): Handbook of mathematical functions with formulas, graphs and mathematical tables, Dover Publications, New York.
- Hastings, C. (1955): Approximations for digital computers, Princeton University Press, Princeton, New Jersey.
- 平井安久・中村 忠 (2006): 巨大数・微小数を処理するための算術システムとその 2 項確率計算への応用. 応用統計学 35(2), 93-111.
- Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994): Continuous univariate distributions, Vol.1, second edition, John Wiley & Sons, New York.
- Molenaar, W. (1973): Approximations to the Poisson, binomial and hypergeometric distribution functions. Mathematisch Centrum.
- 中村 忠・平井安久 (2012): コンピュータ支援による正規確率の 2 項近似アルゴリズム I. Technical Reports of Okayama Statistical Association.
- Patel, J.K. and Read C.B. (1996): Handbook of the normal distribution, Marcel Dekker, New York.
- Shenton, L.R. (1954): Inequalities for the normal integral including a new continued fraction. Biometrika 41, 177-189.
- 山内二郎 (1964): 正規分布に関する近似関数。第 5 回プログラミングシンポジウム報告集, N107-114, 情報処理学会.
- 山内二郎編 (1992): 統計数値表. 日本規格協会.