

変形バケットソートに現れる離散型確率分布と Eulerian 数

土屋 高宏¹ · 中村 永友²

(受付 2008 年 6 月 23 日; 改訂 2008 年 11 月 25 日)

要 旨

バケットソートを変形したソーティング・アルゴリズムに現れる離散型確率分布を導出した。このソーティングの過程で、ある規則性を持つ数が現れる。その数に関係する場合の数を漸化式で表現し、確率分布を与えた。また、確率分布の導出の際に現れる漸化式は Eulerian 数と呼ばれる数列を構成するが、導出過程において Eulerian 数との間に微妙な差異があることを示した。さらに数値実験によりこれを確認した。確率分布についてモーメントなどの基本統計量を求め、正規分布との比較、確率分布近似の精密化を行った。さらに、求めた確率分布と連続型一様分布との関係について言及した。本研究が Eulerian 数を導出する新たな例となることを示した。

キーワード： Eulerian 分布，一様分布，キュムラント，漸近展開，正規分布。

1. はじめに

バケットソートと呼ばれる計算時間が $O(n)$ で高速に行えるソーティング・アルゴリズムがある。このアルゴリズムは並べ替えをしたいデータのとりうる値が k 通りあるとき、あらかじめ k 個のバケツを用意しておき、あるいは動的にバケツを増やしていきながら、各々の数字と対応するバケツにデータを入れていくアルゴリズムである。本論文は、 n 個の連続した数字があり、あらかじめ用意するバケツ数が決められておらず、その数がデータの初期状態に依存する「変形バケットソート」を考え、そのバケツ数がどのような確率分布になるのかを考察する。また、確率分布の導出の際に現れる漸化式は Eulerian 数 (Euler, 1755; Graham et al., 1994) と呼ばれる数列を構成することを示す。また、導出過程において両者間に微妙な差異があるので、これについて言及する。ここで Eulerian 数とは、Eulerian numbers が原語であり、定ベキ化係数、ベキ係数などの訳語もあるようであるが、定まった言い方がないため、この呼び方を用いる。

2 節では問題設定と変形バケットソートのアルゴリズムについて述べ、3 節ではバケツ数の確率分布を導出するための考え方を示して確率分布の漸化式表現を与える。4 節ではバケツ数と Eulerian 数の関係について触れ、導出過程において両者間に微妙な差異があることに言及する。この確率分布について、5 節ではモーメントを求め、6 節では一様分布との関係について述べる。また、7 節では正規分布との比較と確率分布近似の精密化を行う。

¹ 城西大学 理学部：〒350-0295 埼玉県坂戸市けやき台 1-1

² 札幌学院大学 経済学部・総合教育センター：〒069-8555 北海道江別市文京台 11 番地

2. 問題設定

簡単のため数字の書かれたカードを例に説明する. 1から n までの数字が書かれたよくシャッフルされている n 枚のカードが手元にある. カードを小さい順に並べ替えをするために, 次の手順で行う.

- (1) 一番上のカードの数字が k のとき, 数字 $k+1$ のカードがすでにテーブルに置かれていたらその上にカード k を載せる.
- (2) もしテーブル上に $k+1$ が無ければ, どのカードの上にも載せず, テーブルにカード k を置く.
- (3) (1)と(2)を手元のカードが無くなるまで続ける.
- (4) テーブルの上に m 個のカードの束ができる($1 \leq m \leq n$).
- (5) 各束の一番上のカードの数字を小さい順にまとめれば, 並べ替えが完了する.

通常のバケットソートとの相違点は, (1)手元の n 枚のカードが1から n までの重複がないという前提がある点, (2)バケツ数がいくつになるかわからないという点, (3)数字 $k+1$ がすでに置かれていれば k を $k+1$ に載せる点, である. この方法による並べ替えでは, 手順の(4)の段階で最大で n 個, 最小で1個のカードの束ができることになる. このとき, カードをテーブルに置き終わった時点におけるカードの束の数に関する確率分布がどのようなものになるのかを議論する. 「束の数」とはバケットソートにおけるバケツ数に対応する.

数字(カード)の並びが与えられたとき, 変形バケットソートにおけるバケツ数(束の数)は以下のアルゴリズムで直接求めることができる.

アルゴリズム. n 個の数字(カード)の並びを $\nu_1, \nu_2, \dots, \nu_j, \dots, \nu_n$ と表す. このとき, $k < j$ ($k=1, \dots, j-1; j=2, \dots, n$)でかつ $\nu_k = \nu_j + 1$ を満足する総数を

$$m = \#\{k \mid \nu_k = \nu_j + 1, k < j\}$$

とすると, $n - m$ がバケツ数である. ここで, 記号 $\#\{\cdot\}$ は, 集合 $\{\cdot\}$ の要素の数を表す.

解説. n 個の数字(カード)の並びを $\nu_1, \nu_2, \dots, \nu_j, \dots, \nu_n$ として, バケツを n 個用意しておく. j を固定して, $\nu_k = \nu_j + 1$ を満足する ν_k ($k=1, \dots, j-1$)があったときには, ν_k の入ったバケツに ν_j を入れることになるので, バケツ数が1個減ることになる. これを $j=2, \dots, n$ で動かしたとき, $\nu_k = \nu_j + 1$ を満足する総数が m 個あったとする. このとき m 個分だけバケツ数が減ることになるので, 最終的なバケツ数は $n - m$ となる.

3. 束の数の確率分布

本節では関心のある確率分布の導出の基本的な考え方を示し, それに基づいて漸化式により確率分布を表現する.

3.1 束の数

カードの枚数 n が2, 3, 4の場合にカードの出方と束の数を列挙すると以下のようなになる.

$$n=2: (1,2) \cdots 2 \quad (2,1) \cdots 1$$

$$n=3: \begin{cases} (1,2,3) \cdots 3 & (1,3,2) \cdots 2 & (2,1,3) \cdots 2 \\ (2,3,1) \cdots 2 & (3,1,2) \cdots 2 & (3,2,1) \cdots 1 \end{cases}$$

$$n=4: \begin{cases} (1,2,3,4)\cdots 4 & (2,1,3,4)\cdots 3 & (3,1,2,4)\cdots 3 & (4,1,2,3)\cdots 3 \\ (1,2,4,3)\cdots 3 & (2,1,4,3)\cdots 2 & (3,1,4,2)\cdots 3 & (4,1,3,2)\cdots 2 \\ (1,3,2,4)\cdots 3 & (2,3,1,4)\cdots 3 & (3,2,1,4)\cdots 2 & (4,2,1,3)\cdots 2 \\ (1,3,4,2)\cdots 3 & (2,3,4,1)\cdots 3 & (3,2,4,1)\cdots 2 & (4,2,3,1)\cdots 2 \\ (1,4,2,3)\cdots 3 & (2,4,1,3)\cdots 2 & (3,4,1,2)\cdots 3 & (4,3,1,2)\cdots 2 \\ (1,4,3,2)\cdots 2 & (2,4,3,1)\cdots 2 & (3,4,2,1)\cdots 2 & (4,3,2,1)\cdots 1 \end{cases}$$

カッコ内はカードの並びを、その右に束の数を示す。なお、カードはカッコ内の左から順にカードが現れるものとする。例えば、 $n=3$ のとき束の数が2になる場合は4通りあり、 $n=4$ のとき束の数が2または3になる場合はそれぞれ11通りあることがわかる。ちなみに、 n 枚のカードがあるとき、カードの並べ方の総数は $n!$ 通りある。

次に一般の場合について考える。 n 枚のカードの束の数を求めるためには、 $n-1$ 枚のカードの並びにおける束の数を元に検証していく。例えばカードが4枚、つまり $n=4$ の束の数は、3枚のカードの並びに4のカードを1枚追加したときの束の数の変化を考えればよい。3枚のカードの並びの1つに $(1,3,2)$ があるが、これは束の数が2になる。これにカードを1枚追加したとき、4のカードは $(a\ 1\ b\ 3\ c\ 2\ d)$ の4つの a, b, c, d のいずれかの位置に入る。このとき、4が3の左側(a または b)に入るときは束の数は2のままで変化はない。一方、右側(c または d)に入るとき、束の数は1増加して3になる。

以上の考え方から一般に n 枚のカードにおける束の数に関して、次のことが言える。 $n-1$ 枚のカードに数字 n を追加したとき、数字 n が数字 $n-1$ の左側に入るとき束の数は変わらず、右側に入るとき束の数は1増加する。したがって、 n 枚のカードにおける束の数が i になる総数 $M_n(i)$ は、 $n-1$ 枚のカードにおいて、束の数が変わらない場合の数と、束の数が1増加して i になる場合の数の和として得られる。実際、次の漸化式が成り立つ。

$$(3.1) \quad \begin{cases} M_n(1) = M_n(n) = 1 & (n \geq 1), \\ M_n(i) = iM_{n-1}(i) + \{n - (i - 1)\}M_{n-1}(i - 1) & (n \geq 3, i = 2, \dots, n - 1). \end{cases}$$

表1は、 $n=12$ までの束の数 $M_n(i)$ の数表である ($n=13, 14, 15$ は付録Aの表4~5の最下行を参照)。以後、前後の文脈から明らかなきときは、「束の数による分類の場合の数」を単に「束の数」と呼ぶことにする。

漸化式 (3.1) は次節で述べる Eulerian 数を求める式と数学的に同値であり、これは 3.3 節で

表1. 束の数 $M_n(i)$ ($n=1, 2, \dots, 12$).

$n \setminus i$	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	1	1										
3	1	4	1									
4	1	11	11	1								
5	1	26	66	26	1							
6	1	57	302	302	57	1						
7	1	120	1191	2416	1191	120	1					
8	1	247	4293	15619	15619	4293	247	1				
9	1	502	14608	88234	156190	88234	14608	502	1			
10	1	1013	47840	455192	1310354	1310354	455192	47840	1013	1		
11	1	2036	152637	2203488	9738114	15724248	9738114	2203488	152637	2036		
12	1	4083	478271	10187685	66318474	162512286	162512286	66318474	10187685	478271	4083	1

示す.

3.2 Eulerian 数

Eulerian 数は Euler (1755) によって提唱され, 次のように定義されている (Knuth, 1998 など).

$$(3.2) \quad A(n, k) = \sum_{j=0}^k (-1)^j \binom{n+1}{j} (k+1-j)^n \quad (n \geq 0, k \geq 0).$$

また, 漸化式による表現は,

$$(3.3) \quad \begin{cases} A(n, k) = (n-k)A(n-1, k-1) + (k+1)A(n-1, k), & (n \geq 1, n-1 \geq k \geq 1), \\ A(n, 0) = 1 \quad (n \geq 1), \quad A(n, k) = 0 \quad (k \geq n \geq 1) \end{cases}$$

である. この漸化式の証明は右辺に (3.2) 式を代入することにより示すことができる.

Eulerian 数は一般的に知られている Euler 数とは別のものである. これは主に代数学や離散数学の分野で研究されており, 以下のように第 2 種スターリング数 $S(n, k)$ やベルヌーイ数 B_n との関係が知られている (例えば, Graham et al., 1994).

$$A(n, k) = \sum_{j=1}^{n-k} S(n, j) (-1)^{n-j-k} j! \binom{n-j}{k},$$

$$\sum_{k=0}^{n-1} (-1)^k A(n, k) = \frac{2^{n+1}(2^{n+1}-1)B_{n+1}}{n+1}.$$

Eulerian 数に対する確率分布の適用例としては, Kimber (1987) が時系列に関する差の符号検定や壺のモデルに適用している.

3.3 束の数と Eulerian 数の関係

束の数 $M_n(i)$ と $A(n, k)$ の 2 つの数の間には, $k = i - 1$ とおくと, $M_n(i) = A(n, i - 1)$ という関係が成立している. 例えば, $M_6(1) = A(6, 0)$, $M_6(2) = A(6, 1)$, $M_6(3) = A(6, 2), \dots$ である. このように束の数と Eulerian 数が同値であることを通して, (3.1) 式が正しいことを以下に示す.

命題 3.1. 束の数と Eulerian 数の間に $M_n(i) = A(n, i - 1) (i = 1, 2, \dots, n)$ が成り立つ.

証明. 命題 3.1 を示すためには, $\{1, 2, \dots, n\}$ の置換全体の集合を, 束の数によって分類するときの場合の数が, Eulerian 数に等しいことを証明すればよい.

自然数 n に対し, $\Omega_n = \{1, 2, \dots, n\}$ とし, Ω_n 上の置換全体の集合を S_n とする.

$$\tau = \begin{pmatrix} 1 & 2 & \cdots & n \\ \tau(1) & \tau(2) & \cdots & \tau(n) \end{pmatrix} \in S_n$$

に対して, $a(\tau) = \#\{j \mid \tau(j) < \tau(j+1), j = 1, 2, \dots, n-1\}$ を τ の上昇点の数とする. Graham et al. (1994) では, Eulerian 数を

$$A(n, k) = \#\{\tau \in S_n \mid a(\tau) = k\}$$

と定義している. $\tau \in S_n$ に対して, τ の逆置換は,

$$\tau^{-1} = \begin{pmatrix} \tau(1) & \tau(2) & \cdots & \tau(n) \\ 1 & 2 & \cdots & n \end{pmatrix}$$

となる. このとき, $\tau \in S_n$ に対して, $\tau(1), \tau(2), \dots, \tau(n)$ の順に Ω_n の要素が現れるときの束の数を $b(\tau)$ と表すと,

$$(3.4) \quad b(\tau) = 1 + a(\tau^{-1})$$

が成り立つ. これは以下のように証明することができる. $j = 1, 2, \dots, n-1$ について, 命題 P_j を「 $\tau(\alpha) = j, \tau(\beta) = j+1$ となる α と β について, $\alpha < \beta$ 」と定義する. Ω_{k-1} 上の置換に k ($k = 2, 3, \dots, n$) を挿入して Ω_k 上の置換を構成する方法を繰り返して $\tau \in S_n$ を作ることを考える. このとき, 最初に数字 1 のみが与えられると束の数は 1 である. j を増やしていき, 各 P_j が成立するたびに束の数が 1 増加するので,

$$b(\tau) = 1 + \sum_{j=1}^{n-1} I(P_j)$$

と表すことができる. ここで,

$$I(P_j) = \begin{cases} 1 & (P_j \text{ が真}) \\ 0 & (P_j \text{ が偽}) \end{cases}$$

である. 命題 P_j は τ^{-1} を用いて表すと, 「 $\tau^{-1}(j) < \tau^{-1}(j+1)$ 」となるので, $b(\tau) = 1 + \#\{j \mid \tau^{-1}(j) < \tau^{-1}(j+1)\}$ となり, (3.4) 式が成り立つ.

S_n から S_n への写像 $\varphi: \tau \mapsto \tau^{-1}$ は全単射である. したがって, S_n の任意の部分集合 C を写像 φ で変換すると, $\#C = \#\varphi(C)$ であるから,

$$M_n(i) = \#\{\tau \in S_n \mid b(\tau) = i\} = \#\{\tau \in S_n \mid b(\tau^{-1}) = i\}$$

となる. ここで, (3.4) 式より $\{\tau \in S_n \mid b(\tau^{-1}) = i\} = \{\tau \in S_n \mid a(\tau) = i-1\}$ となる. また, Eulerian 数の定義から $\#\{\tau \in S_n \mid a(\tau) = i-1\} = A(n, i-1)$ となるので, $M_n(i) = A(n, i-1)$ が成り立つ. \square

命題 3.1 と Eulerian 数の漸化式 (3.3) から, 漸化式 (3.1) はすべての自然数 n について成り立つことがわかる.

3.4 確率分布

n 枚のカードが i 束になる確率 $P_n(i)$ は, $P_n(i) = M_n(i)/n!$ で定義され, 次の漸化式で与えられる.

束の数の分布 (Eulerian 分布)

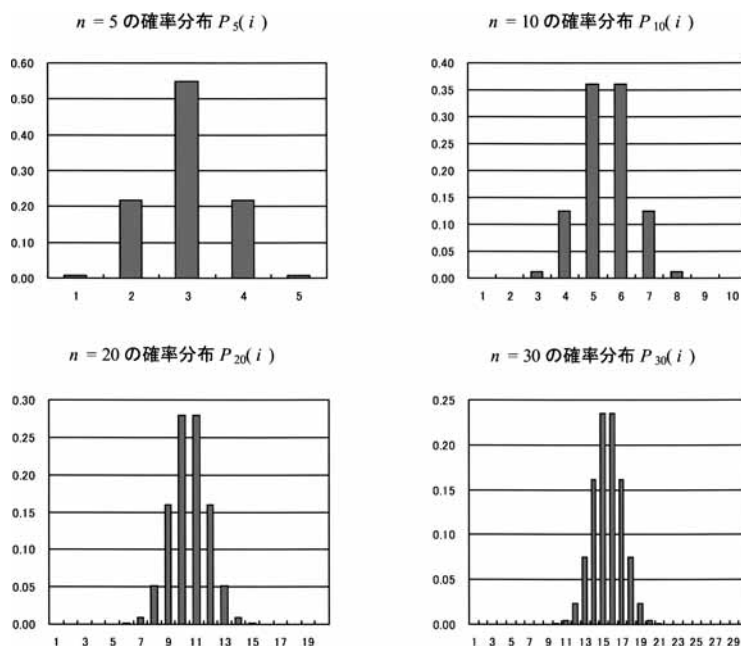
$$(3.5) \quad \begin{cases} P_n(1) = P_n(n) = \frac{1}{n!} & (n \geq 1), \\ P_n(i) = \frac{i}{n} P_{n-1}(i) + \frac{n-(i-1)}{n} P_{n-1}(i-1) & (n \geq 3, i = 2, \dots, n-1). \end{cases}$$

図 1 に束の数 $M_n(i)$ の確率分布 $P_n(i)$ の例を示す.

4. 束の数と Eulerian 数のずれ

表 1 で示した束の数に関する漸化式により与えられる数列は Eulerian 数を構成する. 本節ではこの数列と束の数との関係を考察する.

Eulerian 数とは 1 から n の順列に対して, 上昇点の数 (あるいは, それと同値な増加列・減少列の数) を表すものである. また, Stanley (1997) は下降点の数による分類として Eulerian 数を定義している. 上昇点の数とは n 個の数の並びにおいて, 隣り合う 2 数の大小関係が $<$ と

図 1. $P_n(i)$ ($n=5, 10, 20, 30$) の確率分布.

なる個数 m を表している. 例えば, $(1, 2, 3, 4)$ の並びは $1 < 2 < 3 < 4$ であるから, 上昇点の数は 3 となり, $(3, 2, 4, 1)$ の並びは $3 > 2 < 4 > 1$ であるから, 上昇点の数は 1 となる.

一方, 束の数は m を用いて $m+1$ と表され,

$$\text{束の数} = \text{“上昇点の数”} + 1$$

の関係が成り立っているように見える. しかしながら, 両者の間に食い違いが生じる場合があり, すべての並びについてこの等式が成り立つとは限らない. 例えば, 表 2 のように, 番号 1 と 2 の数字の並びでは両者が一致するが, 番号 3 と 4 では“ずれ”が生じていることがわかる.

$n=4$ のときにずれが生じるのは, この 2 組のみであり, 束の数が 2 と 3 になる場合の総数は「“上昇点の数” + 1」と等しくなる. したがって, 束の数は Eulerian 数を構成していることがわかる.

表 3 は $n=4, 5$ のときの「束の数」と「“上昇点の数” + 1」の一致と不一致を表すクロス表である. $n=4$ の非対角要素, 2 行 3 列と 3 行 2 列の 1 はそれぞれ「束の数」と「“上昇点の数” + 1」が 2 と 3 になるずれと 3 と 2 になるずれが 1 組ずつあることを示している. また, $n=5$

表 2. 束の数と上昇点の数.

番号	数字の並び	束の数	“上昇点の数” + 1	ずれ
1	(2, 1, 4, 3)	2	2	なし
2	(1, 3, 4, 2)	3	3	なし
3	(2, 4, 1, 3)	2	3	あり
4	(3, 1, 4, 2)	3	2	あり

表 3. $n=4,5$ の「束の数」と「上昇点の数+1」の一致と不一致(ずれ).

$n=4$	1	2	3	4	$n=5$	1	2	3	4	5
1	1	0	0	0	1	1	0	0	0	0
2	0	10	1	0	2	0	20	6	0	0
3	0	1	10	0	3	0	6	54	6	0
4	0	0	0	1	4	0	0	6	20	0
和	1	11	11	1	5	0	0	0	0	1
					和	1	26	66	26	1

のときは2と3および3と2になるずれがそれぞれ6組あり, さらに3と4および4と3になるずれもそれぞれ6組あることがわかる. 結果的に行と列の合計が等しくなり, ずれが相殺され, 束の数が Eulerian 数になっている. こうしたずれが生じる理論的根拠は以下のように説明することができる.

$\tau \in S_n$ が与えられたとき, 「上昇点の数+1」は3.3節の記号を用いると, $a(\tau)+1$ であるが, 束の数は $a(\tau)+1$ ではなく, $a(\tau^{-1})+1$ であるから, 両者の値が異なるような τ については, ずれが生じる. 表2の番号3と4の置換は互いに逆置換になっていて, しかも $a(\tau) \neq a(\tau^{-1})$ となっているから, ずれが生じることになる. 逆置換を与える変換 φ は S_n 上の全単射であるから, 対応する場合の数は同じになる.

数学的には束の数と Eulerian 数は同値であるので, ずれの議論は数学的に実りがないと思われるが, 導出過程の相違を指摘したので, そのずれの様子を調べた. $n=6, \dots, 15$ の一致・不一致の様子を付録Aに示す. 表から, ずれの様子が主対角線対称および反主対角線対称であることがわかる. また, ずれの程度は主対角線要素の総和を $n!$ で割った一致率で測ることができる(付録Aの表5). 表を見る限り, n が大きくなるほど一致率は単調に減少している. 導出過程に相違があるが, 先の命題で束の数が Eulerian 数になることが示されたことは, 新たな Eulerian 数を導く一例を示したことになる.

以上の議論より, (3.5)式で与えられる束の数の確率分布 $P_n(i)$ をこれ以降 Eulerian 分布と呼ぶことにする. Eulerian 分布は Eulerian 数を用いて, $P_n(i) = A(n, i-1)/n!$ と表される.

5. 確率分布のモーメント

本節では確率分布の期待値や分散をはじめとするモーメントを与える. Eulerian 分布の平均まわりのモーメントの漸化式表現は Mann (1945) により導出されているが, ここでは Eulerian 分布に対して原点まわりのモーメントの漸化式表現を与える. この漸化式から期待値や分散の一般項を導出する.

定理 5.1. 確率分布(3.5)にしたがう確率変数 X の r 次モーメント $E_n(X^r)$ は, 以下の漸化式で与えられる.

$$(5.1) \quad E_1(X^r) = 1, \quad E_n(X^r) = \sum_{k=0}^r \left\{ \frac{n+1}{n} \binom{r}{k} - \frac{1}{n} \binom{r+1}{k} \right\} E_{n-1}(X^k) \quad (n \geq 2).$$

証明. モーメントの定義より, $n=1$ のとき $E_1(X^r) = 1^r P_1(1) = 1$ である. $n \geq 3$ のとき(3.5)式の i を x に置き換え, そして x^r をかけて $x(x=1, 2, \dots, n)$ について和をとると, 次式が成り

立つ.

$$(5.2) \quad E_n(X^r) = \frac{1}{n}E_{n-1}(X^{r+1}) + \frac{1}{n}E_{n-1}[(n-X)(X+1)^r].$$

(5.2) 式の右辺第 2 項は

$$\begin{aligned} & \frac{1}{n}E_{n-1}[(n-X)(X+1)^r] \\ &= E_{n-1}[(X+1)^r] - \frac{1}{n}E_{n-1}[X(X+1)^r] \\ &= \sum_{k=0}^r \binom{r}{k} E_{n-1}(X^k) - \frac{1}{n} \sum_{k=0}^{r-1} \binom{r}{k} E_{n-1}(X^{k+1}) - \frac{1}{n}E_{n-1}(X^{r+1}) \end{aligned}$$

と変形できるので, (5.2) 式は

$$\begin{aligned} (5.3) \quad E_n(X^r) &= \sum_{k=0}^r \binom{r}{k} E_{n-1}(X^k) - \frac{1}{n} \sum_{k=0}^{r-1} \binom{r}{k} E_{n-1}(X^{k+1}) \\ &= \sum_{k=0}^r \binom{r}{k} E_{n-1}(X^k) - \frac{1}{n} \left\{ \sum_{k=0}^r \binom{r+1}{k} E_{n-1}(X^k) - \sum_{k=0}^r \binom{r}{k} E_{n-1}(X^k) \right\} \\ &= \sum_{k=0}^r \left\{ \frac{n+1}{n} \binom{r}{k} - \frac{1}{n} \binom{r+1}{k} \right\} E_{n-1}(X^k) \end{aligned}$$

となる. $n=2$ のとき (3.5) 式から

$$(5.4) \quad E_2(X^r) = 1^r P_2(1) + 2^r P(2) = \frac{1}{2}(1+2^r)$$

を得る. ここで, (5.3) 式で $n=2$ とおいた式は (5.4) 式と一致することがわかる. したがって (5.3) 式は $n \geq 2$ のときにも成り立つ. \square

この証明は Mann (1945) に基づく方法である. また, 土屋・中村 (2007) には離散型確率分布の期待値の定義式による証明がある.

次に, Eulerian 分布に対する期待値と分散は次の定理により得られる.

定理 5.2. 確率分布 (3.5) にしたがう確率変数 X の期待値 $E_n(X)$ と分散 $V_n(X)$ は, それぞれ

$$\begin{aligned} E_n(X) &= \frac{n+1}{2} \quad (n \geq 1), \\ V_1(X) &= 0, \quad V_n(X) = \frac{n+1}{12} \quad (n \geq 2) \end{aligned}$$

で与えられる.

証明. $r=1$ のとき (5.1) 式は

$$(5.5) \quad E_1(X) = 1, \quad E_n(X) = \frac{n-1}{n}E_{n-1}(X) + 1$$

となり, さらに $B_n(X) = nE_n(X)$ とおくと (5.5) 式は

$$(5.6) \quad B_1(X) = E_1(X) = 1, \quad B_n(X) = B_{n-1}(X) + n$$

となる. (5.6) 式から $n \geq 2$ に対して,

$$(5.7) \quad B_n(X) = B_1(X) + \sum_{k=1}^{n-1} k = \frac{n(n+1)}{2}$$

を得る. (5.7)式は $n=1$ のときも成り立つ. よって, 期待値 $E_n(X)$ の一般項は

$$E_n(X) = \frac{B_n(X)}{n} = \frac{n+1}{2} \quad (n \geq 1)$$

で与えられる.

$r=2$ のとき (5.1)式は

$$E_1(X^2) = 1, \quad E_n(X^2) = \frac{n-2}{n}E_{n-1}(X^2) + \frac{2n-1}{n}E_{n-1}(X) + 1$$

と表される. よって, 分散 $V_n(X)$ に関して

$$V_1(X) = E_1(X^2) - \{E_1(X)\}^2 = 0,$$

$$(5.8) \quad V_n(X) = E_n(X^2) - \{E_n(X)\}^2 = \frac{n-2}{n}V_{n-1}(X) + \frac{1}{4}$$

を得る. (5.8)式が

$$(5.9) \quad V_n(X) - (an + b) = \frac{n-2}{n}[V_{n-1}(X) - \{a(n-1) + b\}]$$

を満たすように定数 a, b の値を求めると, $a = b = 1/12$ である. ここで, $C_n(X) = V_n(X) - \{(n+1)/12\}$ とおくと, (5.9)式は $C_n(X) = \{(n-2)/n\}C_{n-1}(X)$ となるので, $n \geq 2$ に対して $C_n(X) = 0$ を得る. したがって,

$$V_n(X) = \frac{n+1}{12} \quad (n \geq 2)$$

となる. \square

6. Eulerian 分布と一様分布

Eulerian 分布は一様分布と以下のような関連がある. Hensley (1982) は, Eulerian 分布が次の確率と等しいことをラプラス変換を用いて示した.

$$(6.1) \quad P_n(k+1) = \Pr\left(\sum_{j=1}^n X_j \in [k, k+1]\right).$$

ここで X_j は $(0,1)$ 上の一様分布にしたがう確率変数である. ここでは, このことを区間 $[k, k+1]$ で直接積分することにより示す.

一様分布にしたがう確率変数の和の確率密度関数は次のように表される (Mood et al., 1974, p.238).

$$(6.2) \quad f(s) = \sum_{j=0}^{n-1} \frac{1}{(n-1)!} \left[s^{n-1} - \binom{n}{1}(s-1)^{n-1} + \binom{n}{2}(s-2)^{n-1} \right. \\ \left. - \dots + (-1)^j \binom{n}{j}(s-j)^{n-1} \right] I_{(j,j+1]}(s).$$

ただし,

$$I_R(s) = \begin{cases} 1 & (s \in R) \\ 0 & (s \notin R) \end{cases}$$

である.

(6.2)式の積分は、区間 $[k, k+1]$ に対応する区間 $(j, j+1]$ のみで考えればよいから、和の記号を除いて j を k で置きかえると以下のようなになる。

$$\begin{aligned}
 & \int_k^{k+1} f(s) ds \\
 &= \frac{1}{n!} \left[s^n - \binom{n}{1} (s-1)^n + \binom{n}{2} (s-2)^n - \cdots + (-1)^k \binom{n}{k} (s-k)^n \right]_k^{k+1} \\
 &= \frac{1}{n!} \left[(k+1)^n - \binom{n}{1} k^n + \binom{n}{2} (k-1)^n - \cdots + (-1)^k \binom{n}{k} \right] \\
 &\quad - \left[k^n - \binom{n}{1} (k-1)^n + \binom{n}{2} (k-2)^n - \cdots + (-1)^k \binom{n}{k} \right] \\
 &= \frac{1}{n!} \left[(k+1)^n - \binom{n+1}{1} k^n + \binom{n+1}{2} (k-1)^n - \cdots + (-1)^k \binom{n+1}{k} \right] \\
 &= \frac{1}{n!} \sum_{j=0}^k (-1)^j \binom{n+1}{j} (k+1-j)^n = \frac{1}{n!} A(n, k) = P_n(k+1).
 \end{aligned}$$

したがって、(6.1)式が証明できた。

また、同時に次の別の結果を示すことができる。(6.2)式において n の代わりに $n+1$ とおき、 $s=k+1$ とおくと、

$$\begin{aligned}
 f(s) &= \frac{1}{n!} \left[s^n - \binom{n+1}{1} (s-1)^n + \binom{n+1}{2} (s-2)^n - \cdots + (-1)^k \binom{n+1}{k} (s-k)^n \right] \\
 &= \frac{1}{n!} \sum_{j=0}^k (-1)^j \binom{n+1}{j} (k+1-j)^n = P_n(k+1)
 \end{aligned}$$

となり、Eulerian 分布は $n+1$ 個の連続型一様確率変数の和の確率密度関数と一致する。つまりこれは、離散型確率分布の確率と連続型確率分布の確率密度という性質の違う2つの一致を意味し、大変興味深い結果である。

7. 確率分布の近似

Eulerian 分布の正規分布への近似については、Mann (1945) や Hensley (1982) によって研究されている。Mann は Eulerian 数 $A(n, k)$ のモーメントに関する漸化式から、 $A(n, k)$ を基準化した変数が $n \rightarrow \infty$ のとき、標準正規分布のモーメントに収束することを示した。また、Hensley は Eulerian 分布と一様確率変数の和の分布の関係を利用して、中心極限定理により正規分布への近似を行った。ここでは、高次のキュムラントを使って確率分布近似の精密化を行う。

7.1 近似の精密化

モーメントに関する漸化式から、より高次のキュムラントを導出することができる。6次までのキュムラント $\kappa_r^{(n)}$ ($r=1, 2, \dots, 6$) は次の定理で与えられる。

定理 7.1. 確率分布 (3.5) の 6 次までのキュムラント $\kappa_r^{(n)}$ ($r=1, 2, \dots, 6$) は,

$$\begin{aligned} \kappa_1^{(n)} &= \frac{n+1}{2} \quad (n \geq 1), & \kappa_2^{(n)} &= \frac{n+1}{12} \quad (n \geq 2), & \kappa_3^{(n)} &= 0 \quad (n \geq 1), \\ \kappa_4^{(n)} &= -\frac{n+1}{120} \quad (n \geq 4), & \kappa_5^{(n)} &= 0 \quad (n \geq 1), & \kappa_6^{(n)} &= \frac{n+1}{252} \quad (n \geq 6) \end{aligned}$$

である.

証明. 付録 C 参照. □

確率分布 (3.5) は対称分布であるから, 平均を除く奇数次のキュムラントは 0 となる. また, 1 次 of キュムラントは $(n+1)(1+B_1)$, 偶数次のキュムラントは $(n+1)B_r/r$ ($r \leq 6$) として表すことができる. 8 次以上のキュムラントは, その式が非常に複雑であるためここでは示さない. ここで, B_r は次の関係式を満たすベルヌーイ数である.

$$\frac{t}{e^t - 1} = \sum_{r=0}^{\infty} \frac{B_r}{r!} t^r.$$

$r=6$ までのベルヌーイ数 B_r は以下の通りである.

r	1	2	3	4	5	6
B_r	$-\frac{1}{2}$	$\frac{1}{6}$	0	$-\frac{1}{30}$	0	$\frac{1}{42}$

これは, $n+1$ 個の一樣分布にしたがう確率変数の和の分布のキュムラントと酷似している. 違いは, 一樣分布は任意の自然数 n について成り立つが, Eulerian 分布は n がキュムラントの次数以上で成り立つという点である. n がキュムラントの次数より小さいときは一樣分布のキュムラントと異なる.

したがって, 確率分布 (3.5) の r 次 of キュムラント $\kappa_r^{(n)}$ ($r \leq 6$) は, $(0,1)$ 上の一樣分布にしたがう $n+1$ 個の互いに独立で同一な確率変数の和のキュムラントと一致する.

次に, 確率分布の精密化は以下の定理により得られる.

定理 7.2. T_n を確率分布 (3.5) にしたがう確率変数とすると,

$$(7.1) \quad X_n = \frac{\sqrt{N}(T_n - N/2)}{N/\sqrt{12}} = \sqrt{\frac{N}{12}} \left(\frac{T_n}{N} - \frac{1}{2} \right)$$

とおくと, 確率関数 $f(x) = \Pr(X_n = x)$ ($n \geq 6$) に対して, 次のような $1/N^2$ の項までの漸近展開式を得る.

$$f(x) \approx \phi(x) \left\{ 1 - \frac{1}{N} a_1(x) + \frac{1}{N^2} a_2(x) \right\}.$$

ここで, x のとりうる値は $T_n = i$ ($i=1, 2, \dots, n$) を (7.1) 式に代入した

$$(7.2) \quad x = \sqrt{\frac{N}{12}} \left(\frac{i}{N} - \frac{1}{2} \right)$$

で与えられる. また, $N = n+1$, $\phi(x)$ は標準正規密度関数, 係数 $a_1(x)$, $a_2(x)$ は,

$$a_1(x) = \frac{1}{20} H_4(x), \quad a_2(x) = \frac{1}{105} H_6(x) + \frac{1}{800} H_8(x)$$

である. ただし, $H_j(x)$ ($j=1, 2, \dots$) は j 次 of エルミート多項式である.

証明. $N = n+1$, $\mu = N/2$, $\sigma = N/\sqrt{12}$ とすると,

$$X_n = \frac{\sqrt{N}(T_n - \mu)}{\sigma}$$

は漸近的に標準正規分布 $N(0, 1)$ に収束する. このとき, X_n の特性関数は次の式で与えられる.

$$(7.3) \quad \begin{aligned} \psi_{X_n}(t) &= E(e^{itX_n}) = \exp\left(-\frac{\sqrt{N}\mu}{\sigma}it\right) E\left[\exp\left(\frac{\sqrt{N}T_n}{\sigma}it\right)\right] \\ &= \exp\left(-\frac{\sqrt{N}\mu}{\sigma}it\right) \exp\left[\left(\frac{\sqrt{N}}{\sigma}it\right)\kappa_1^{(n)} + \frac{1}{2}\left(\frac{\sqrt{N}}{\sigma}it\right)^2\kappa_2^{(n)}\right. \\ &\quad \left. + \frac{1}{3!}\left(\frac{\sqrt{N}}{\sigma}it\right)^3\kappa_3^{(n)} + \frac{1}{4!}\left(\frac{\sqrt{N}}{\sigma}it\right)^4\kappa_4^{(n)} + \dots\right]. \end{aligned}$$

ここで, $\kappa_r^{(n)}$ は r 次のキュムラントである. さらに (7.3) 式に $\kappa_1^{(n)} = \mu = N/2$, $\kappa_2^{(n)} = \sigma^2 = N/12$, $\kappa_3^{(n)} = 0$, $\kappa_4^{(n)} = -N/120$, $\kappa_5^{(n)} = 0$, $\kappa_6^{(n)} = N/252$ を代入すると,

$$(7.4) \quad \psi_{X_n}(t) = \exp\left(-\frac{t^2}{2}\right) \left[1 - \frac{1}{N} \left\{\frac{1}{20}(it)^4\right\} + \frac{1}{N^2} \left\{\frac{1}{105}(it)^6 + \frac{1}{800}(it)^8\right\} + \dots\right]$$

となる. いま,

$$\frac{1}{2\pi} \int (it)^r e^{-t^2/2} e^{-itx} dx = H_r(x) \phi(x)$$

を用いて (7.4) 式を反転すると, $\Pr(X_n = x)$ に対して $1/N^2$ の項までの式

$$f(x) \approx \phi(x) \left[1 - \frac{1}{N} \left\{\frac{1}{20}H_4(x)\right\} + \frac{1}{N^2} \left\{\frac{1}{105}H_6(x) + \frac{1}{800}H_8(x)\right\}\right]$$

を得る. □

7.2 数値近似

図 2 に確率分布 (3.5) の近似誤差として, 真の値の差 E と $\log_{10}|E|$ を示す ($n=6, 10, 20$). 次式の第 1 項 (正規近似), $1/N$ までの項, $1/N^2$ までの項の近似式がそれぞれ R, Z1, Z2 である.

$$(7.5) \quad \begin{aligned} \Pr(T_n = i) &\approx \frac{1}{\sqrt{2\pi N/12}} e^{-\frac{(i-N/2)^2}{N/6}} \left[1 - \frac{1}{N} \left\{\frac{1}{20}H_4(x)\right\}\right. \\ &\quad \left. + \frac{1}{N^2} \left\{\frac{1}{105}H_6(x) + \frac{1}{800}H_8(x)\right\}\right]. \end{aligned}$$

ここで, $i=1, 2, \dots, n$ であり, x は (7.2) 式で与えられる. 図中の R1 は (6.1) 式 (Hensley, 1982) に基づく正規分布近似である.

図 2 の左図を見ると, 近似式 R でも誤差が ± 0.003 の範囲内となっており, 十分良い近似値を与えているが, 近似式 Z1, Z2 はすべての値について近似精度が改善されていることがわかる. 図 2 の右図では, 提案方法が相対的に近似精度がよいことがわかる.

なお, 離散型確率分布で行うような補正, 例えば二項分布の補正のようなことはしていないが, 提案した確率分布の近似の方法で十分な精度が得られたことを付け加えておく.

8. おわりに

本論文は, 変形バケツソートにおけるバケツ数の離散型確率分布を導出し, その漸化式表現を与えた. 漸化式から導かれる数列は Eulerian 数を構成することを示した. また, 束の数と Eulerian 数の導出過程にずれがあることを示し, 数値的に検証した. 以上より, 束の数は Eulerian 数を導出する新しい例を与えたことになる. また, Eulerian 分布の数学的特性や一様分布との関係を明らかにし, 一様分布にしたがう確率変数の和の分布を利用して, 従来の正規分布近似よりも精密な分布近似を行った.

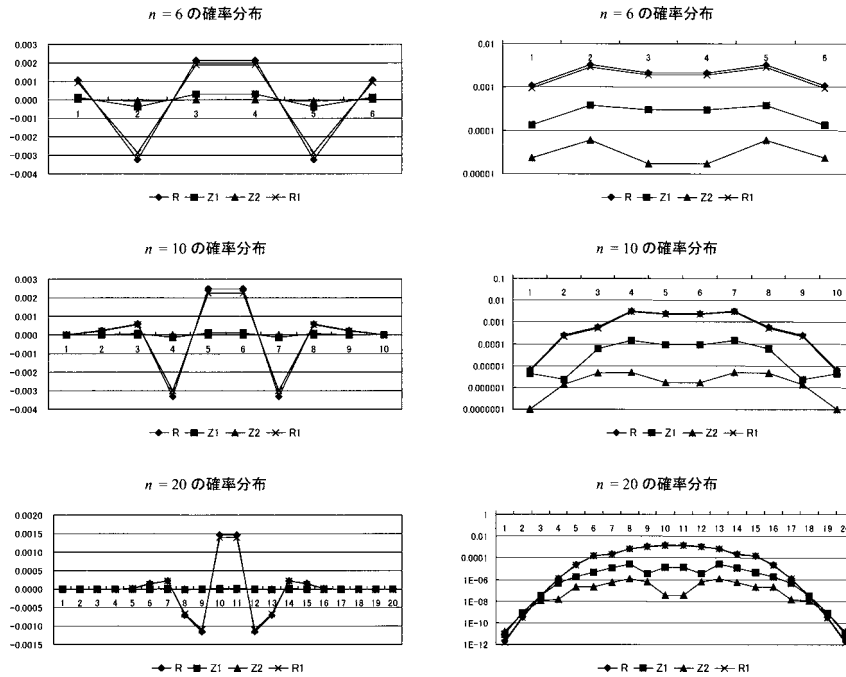


図 2. 正規分布近似と近似の精密化. 左は $n = 6, 10, 20$ の確率分布の値を (7.5) 式の第 1 項 (正規近似), $1/N$ までの項, $1/N^2$ までの項を用いて計算したときの誤差 $E = \text{近似値} - \text{真の値}$ (それぞれ, R, Z1, Z2) と (6.1) 式の正規分布近似による誤差 E . 右は誤差 $E_1 = \log_{10} |\text{近似値} - \text{真の値}|$ である (縦軸は対数軸).

今後の課題としては, (1) 束の数に関する漸化式の直接的な証明, (2) n 枚の「ずれ」を表す各クロス表の要素第 k 表の i 行 j 列が, k に関して別の意味のある数列を構成する可能性があるので, できる限り明らかにすること, などが挙げられる.

謝 辞

本研究を完成するにあたって, 統計数理研究所 平野勝臣名誉教授と九州大学数理学研究院 小西貞則教授に貴重なコメントをいただきました. 匿名査読者には命題 3.1 の証明を示していただくと同時に, 有益な助言と不備の指摘をいただき, 本論文の質を高めることができました. ここに記して御礼を申し上げます. また, 本研究は中村が 2008 年度札幌学院大学長期国内研究員の期間内に完成させました.

付 録

A. 「束の数」と「上昇点の数 + 1」の一致と不一致

$\{1, 2, \dots, n\}$ の n 個の数字 (カード) のすべての並べ替え (組み合わせ) $\{\tau_1, \tau_2, \dots, \tau_n\}$ から得られる「束の数」と「上昇点の数 + 1」の一致と不一致の総数, および一致率をまとめたものを表 4 と表 5 にそれぞれ示す. 任意の順列が与えられたときに, これらが一致するの否かを一

瞬で判断できる理論がないため、すべてについて確認した。行和および列和は Eulerian 数になる。また、主対角線に対して対称なので、行と列は「束の数」と「“上昇点の数”+1」のどちらでも良い。

B. 束の数と Eulerian 数の計算時間について

ここに示す表 6 は付録 A の表 4 を計算するときにかかった時間である。本論文で計算に用いた計算機の仕様は CPU: Intel^(R) Core^(TM)2 Duo 3.00GHz である。確認した束の数 (Eulerian 数) のオーダー n 、基礎となる計算量の $n!$ 、および実計算時間を示す。 $n=15$ までは計算確認済みであるが、明らかに組み合わせ爆発している $n=16$ は予測時間を示した。

表 4. 束の数と上昇点の数の関係。

$n=6$						$n=8$							
	1	2	3	4	5 6		1	2	3	4	5	6	7 8
1	1	0	0	0	0 0	1	1	0	0	0	0	0	0 0
2	0	35	21	1	0 0	2	0	84	126	36	1	0	0 0
3	0	21	210	70	1 0	3	0	126	1773	1980	405	9	0 0
4	0	1	70	210	21 0	4	0	36	1980	8436	4761	405	1 0
5	0	0	1	21	35 0	5	0	1	405	4761	8436	1980	36 0
6	0	0	0	0	0 1	6	0	0	9	405	1980	1773	126 0
和	1	57	302	302	57 1	7	0	0	0	1	36	126	84 0
						8	0	0	0	0	0	0	0 1
						和	1	247	4293	15619	15619	4293	247 1

$n=7$						$n=9$								
	1	2	3	4	5 6 7		1	2	3	4	5	6	7 8 9	
1	1	0	0	0	0 0 0	1	1	0	0	0	0	0	0 0 0	
2	0	56	56	8	0 0 0	2	0	120	252	120	10	0	0 0 0	
3	0	56	659	440	36 0 0	3	0	252	4245	7150	2750	120	1 0 0	
4	0	8	440	1520	440 8 0	4	0	120	7150	38204	35410	7140	210 0 0	
5	0	0	36	440	659 56 0	5	0	10	2750	35410	79850	35410	2750 10 0	
6	0	0	0	8	56 56 0	6	0	0	210	7140	35410	38204	7150 120 0	
7	0	0	0	0	0 0 1	7	0	0	1	210	2750	7150	4245 252 0	
和	1	120	1191	2416	1191 120 1	8	0	0	0	0	10	120	252 120 0	
						9	0	0	0	0	0	0	0 0 1	
						和	1	502	14608	88234	156190	88234	14608 502 1	

$n=10$										
	1	2	3	4	5	6	7	8	9	10
1	1	0	0	0	0	0	0	0	0	0
2	0	165	462	330	55	1	0	0	0	0
3	0	462	9273	22023	13750	2266	66	0	0	0
4	0	330	22023	147301	203610	75306	6556	66	0	0
5	0	55	13750	203610	592130	423236	75306	2266	1	0
6	0	1	2266	75306	423236	592130	203610	13750	55	0
7	0	0	66	6556	75306	203610	147301	22023	330	0
8	0	0	0	66	2266	13750	22023	92732	462	0
9	0	0	0	0	1	55	330	462	165	0
10	0	0	0	0	0	0	0	0	0	1
和	1	1013	47840	455192	1310354	1310354	455192	47840	1013	1

表 4. つづき

$n = 11$	1	2	3	4	5	6	7	8	9	10	11	
1	1	0	0	0	0	0	0	0	0	0	0	
2	0	220	792	792	220	12	0	0	0	0	0	
3	0	792	18810	60072	55638	16092	1221	12	0	0	0	
4	0	792	60072	498640	965976	572232	101684	4080	12	0	0	
5	0	220	55638	965976	3599619	3797232	1216524	101684	1221	0	0	
6	0	12	16092	572232	3797232	6953112	3797232	572232	16092	12	0	
7	0	0	1221	101684	1216524	3797232	3599619	965976	55638	220	0	
8	0	0	12	4080	101684	572232	965976	498640	60072	792	0	
9	0	0	0	12	1221	16092	55638	60072	18810	792	0	
10	0	0	0	0	0	12	220	792	792	220	0	
11	0	0	0	0	0	0	0	0	0	0	1	
和	1	2036	152637	2203488	9738114	15724248	9738114	2203488	152637	2036	1	
$n = 12$	1	2	3	4	5	6	7	8	9	10	11	12
1	1	0	0	0	0	0	0	0	0	0	0	0
2	0	286	1287	1716	715	78	1	0	0	0	0	0
3	0	1287	35893	148798	192348	86788	12714	442	1	0	0	0
4	0	1716	148798	1516450	3941197	3438240	1042847	96642	1794	1	0	0
5	0	715	192348	3941197	18584241	27330862	13870415	2301612	96642	442	0	0
6	0	78	86788	3438240	27330862	64850149	51880192	13870415	1042847	12714	1	0
7	0	1	12714	1042847	13870415	51880192	64850149	27330862	3438240	86788	78	0
8	0	0	442	96642	2301612	13870415	27330862	18584241	3941197	192348	715	0
9	0	0	1	1794	96642	1042847	3438240	3941197	1516450	148798	1716	0
10	0	0	0	1	442	12714	86788	192348	148798	35893	1287	0
11	0	0	0	0	0	1	78	715	1716	1287	286	0
12	0	0	0	0	0	0	0	0	0	0	1	1
和	1	4083	478271	10187685	66318474	162512286	162512286	66318474	10187685	478271	4083	1

表 4. つづき

$n = 13$	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	1	0	0	0	0	0	0	0	0	0	0	0	0	
2	0	364	2002	3432	2002	364	14	0	0	0	0	0	0	
3	0	2002	65065	340704	588199	383656	92897	7098	105	0	0	0	0	
4	0	3432	340704	4216408	14230748	17266340	8044596	1362452	68210	560	0	0	0	
5	0	2002	588199	14230748	83751213	164901660	122428656	34152132	3158610	68210	105	0	0	
6	0	364	383656	17266340	164901660	499463964	555063978	233019864	34152132	1362452	7098	0	0	
7	0	14	92897	8044596	122428656	555063978	903911722	555063978	122428656	8044596	92897	14	0	
8	0	0	7098	1362452	34152132	233019864	555063978	499463964	164901660	17266340	383656	364	0	
9	0	0	105	68210	3158610	34152132	122428656	164901660	83751213	14230748	588199	2002	0	
10	0	0	0	560	68210	1362452	8044596	17266340	14230748	4216408	340704	3432	0	
11	0	0	0	0	105	7098	92897	383656	588199	340704	65065	2002	0	
12	0	0	0	0	0	0	14	364	2002	3432	2002	364	0	
13	0	0	0	0	0	0	0	0	0	0	0	0	1	
和	1	8178	1479726	45533450	423281535	1505621508	2275172004	1505621508	423281535	45533450	1479726	8178	1	
$n = 14$	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	455	3003	6435	5005	1365	105	1	0	0	0	0	0	0
3	0	3003	112905	730665	1629435	1456560	529956	71940	2835	15	0	0	0	0
4	0	6435	730665	10865585	46433475	75169560	50184540	13633740	1349931	36735	120	0	0	0
5	0	5005	1629435	46433475	336576825	860578230	885230850	375891370	62035485	3324750	36735	15	0	0
6	0	1365	1456560	75169560	860578230	3276306027	4867702245	2973615765	725044860	62035485	1349931	2835	0	0
7	0	105	529956	50184540	885230850	4867702245	10159768150	8644547430	2973615765	375891370	13633740	71940	1	0
8	0	1	71940	13633740	375891370	2973615765	8644547430	10159768150	4867702245	885230850	50184540	529956	105	0
9	0	0	2835	1349931	62035485	725044860	2973615765	4867702245	3276306027	860578230	75169560	1456560	1365	0
10	0	0	15	36735	3324750	62035485	375891370	885230850	860578230	336576825	46433475	1629435	5005	0
11	0	0	0	120	36735	1349931	13633740	50184540	75169560	46433475	10865585	730665	6435	0
12	0	0	0	0	15	2835	71940	529956	1456560	1629435	730665	112905	3003	0
13	0	0	0	0	0	0	1	105	1365	5005	6435	3003	455	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	1
和	1	16369	4537314	198410786	2571742175	12843262863	27971176092	27971176092	12843262863	2571742175	198410786	4537314	16369	1

表 4. つづき

$n = 15$	1	2	3	4	5	6	7	8
1	1	0	0	0	0	0	0	0
2	0	560	4368	11440	11440	4368	560	16
3	0	4368	188682	1482416	4160120	4899264	2511376	536384
4	0	11440	1482416	26232784	139089120	291102560	265085216	106311200
5	0	11440	4160120	139089120	1226726340	3976585056	5447334872	3296123728
6	0	4368	4899264	291102560	3976585056	18757217104	36181108688	30525772272
7	0	560	2511376	265085216	5447334872	36181108688	95223285735	108194712656
8	0	16	536384	106311200	3296123728	30525772272	108194712656	163291904960
9	0	0	41328	17712368	860615124	11256924256	53938266232	108194712656
10	0	0	800	1048560	87791520	1694544480	11256924256	30525772272
11	0	0	1	15232	2776968	87791520	860615124	3296123728
12	0	0	0	16	15232	1048560	17712368	106311200
13	0	0	0	0	1	800	41328	536384
14	0	0	0	0	0	0	0	16
15	0	0	0	0	0	0	0	0
和	1	32752	13824739	848090912	15041229521	102776998928	311387598411	447538817472

$n = 15$	9	10	11	12	13	14	15
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	41328	800	1	0	0	0	0
4	17712368	1048560	15232	16	0	0	0
5	860615124	87791520	2776968	15232	1	0	0
6	11256924256	1694544480	87791520	1048560	800	0	0
7	53938266232	11256924256	860615124	17712368	41328	0	0
8	108194712656	30525772272	3296123728	106311200	536384	16	0
9	95223285735	36181108688	5447334872	265085216	2511376	560	0
10	36181108688	18757217104	3976585056	291102560	4899264	4368	0
11	5447334872	3976585056	1226726340	139089120	4160120	11440	0
12	265085216	291102560	139089120	26232784	1482416	11440	0
13	2511376	4899264	4160120	1482416	188682	4368	0
14	560	4368	11440	11440	4368	560	0
15	0	0	0	0	0	0	1
和	311387598411	102776998928	15041229521	848090912	13824739	32752	1

表 5. 束の数と Eulerian 数の一致率.

n	≤ 3	4	5	6	7	8	9	10	11	12	13	14	15
一致率	1.000	0.917	0.800	0.683	0.586	0.511	0.455	0.426	0.380	0.355	0.334	0.316	0.301

表 6. 付録 A の表を計算するのに要した時間.

n	1	2	3	4	5	6	7	8	9	10	11
$n!$ (計算量)	1	2	6	24	120	720	5040	40320	362880	3628800	39916800
計算時間									< 1sec	1sec	10sec

n	12	13	14	15	16
$n!$ (計算量)	479001600	6227020800	87178291200	1307674368000	20922789888000
計算時間	2.4min	33.2min	8.4hour	5.6day	(3.2month)

C. 定理7.1の証明

C.1 キュムラントとモーメントの関係

キュムラント $\kappa_r^{(n)}$ とモーメント $E_n(X^r)$ の関係

$$\exp\left\{\sum_{r=1}^{\infty} \frac{t^r}{r!} \kappa_r^{(n)}\right\} = \sum_{r=0}^{\infty} \frac{t^r}{r!} E_n(X^r)$$

から、キュムラントとモーメントの $r=6$ までの関係は以下のように与えられる。ただし、 $\kappa_r^{(n)} = \kappa_r$, $E_n(X^r) = \mu'_r$ とする。

$$\begin{aligned} \mu'_1 &= \kappa_1, \\ \mu'_2 &= \kappa_2 + \kappa_1^2, \\ \mu'_3 &= \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3, \\ \mu'_4 &= \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4, \\ \mu'_5 &= \kappa_5 + 5\kappa_4\kappa_1 + 10\kappa_3\kappa_2 + 10\kappa_3\kappa_1^2 + 15\kappa_2^2\kappa_1 + 10\kappa_2\kappa_1^3 + \kappa_1^5, \\ \mu'_6 &= \kappa_6 + 6\kappa_5\kappa_1 + 15\kappa_4\kappa_2 + 15\kappa_4\kappa_1^2 + 10\kappa_3^2 + 60\kappa_3\kappa_2\kappa_1 + 20\kappa_3\kappa_1^3 + 15\kappa_2^3 \\ &\quad + 45\kappa_2^2\kappa_1^2 + 15\kappa_2\kappa_1^4 + \kappa_1^6. \end{aligned}$$

C.2 1次, 2次のキュムラント

定理3.2より明らか。

C.3 奇数次のキュムラント

X の確率分布は対称であるから、 $n \geq 1$ に対する $r \geq 3$ 次のキュムラントは0となる。

C.4 4次のキュムラント

$r=4$ のとき、(5.1)式は

$$\begin{aligned} E_1(X^4) &= 1, \\ E_n(X^4) &= \frac{n-4}{n} E_{n-1}(X^4) + \frac{4n-6}{n} E_{n-1}(X^3) + \frac{6n-4}{n} E_{n-1}(X^2) \\ &\quad + \frac{4n-1}{n} E_{n-1}(X) + 1 \end{aligned}$$

と表される。ここで、3次以下のモーメントおよびキュムラントとモーメントの関係より、 $n \geq 3$ に対して

$$(C.1) \quad \kappa_4^{(n)} = \frac{n-4}{n} \kappa_4^{(n-1)} - \frac{1}{24}$$

を得る(土屋・中村, 2007)。ここで、 $C_n(X) = \kappa_4^{(n)} + \{(n+1)/120\}$ とおくと、(C.1)式は $C_n(X) = \{(n-4)/n\} C_{n-1}(X)$ となるから、 $n \geq 4$ に対して4次のキュムラントは

$$\kappa_4^{(n)} = -\frac{n+1}{120}$$

となる。

C.5 6次のキュムラント

$r=6$ のとき、(5.1)式は

$$E_1(X^6) = 1,$$

$$E_n(X^6) = \frac{n-6}{n}E_{n-1}(X^6) + \frac{6n-15}{n}E_{n-1}(X^5) + \frac{15n-20}{n}E_{n-1}(X^4) \\ + \frac{20n-15}{n}E_{n-1}(X^3) + \frac{15n-6}{n}E_{n-1}(X^2) + \frac{6n-1}{n}E_{n-1}(X) + 1$$

と表される。ここで、5次以下のモーメントおよびキュムラントとモーメントの関係より、 $n \geq 5$ に対して

$$(C.2) \quad \kappa_6^{(n)} = \frac{n-6}{n}\kappa_4^{(n-1)} + \frac{1}{36}$$

を得る(土屋・中村, 2007)。ここで、 $C_n(X) = \kappa_4^{(n)} - \{(n+1)/252\}$ とおくと、(C.2)式は $C_n(X) = \{(n-6)/n\}C_{n-1}(X)$ となるから、 $n \geq 6$ に対して6次のキュムラントは

$$\kappa_6^{(n)} = \frac{n+1}{252}$$

となる。

参 考 文 献

- Euler, L. (1755). *Institutiones Calculi Differentialis*. (John D. Blanton 訳(2000)). *Foundations of Differential Calculus*, Springer, New York)
- Graham, R. L., Knuth, D. E. and Patashnik, O. (1994). Eulerian numbers, *Concrete Mathematics: A Foundation for Computer Science*, 2nd ed., Section 6.2, 267–272, Addison-Wesley, Boston.
- Hensley, D. (1982). Eulerian numbers and the unit cube, *Fibonacci Quarterly*, **20**, 344–348.
- Kimber, A. C. (1987). Eulerian numbers and links with some statistical procedures, *Utilitas Mathematica*, **31**, 57–65.
- Knuth, D. E. (1998). Runs, *The Art of Computer Programming*, 2nd ed., Section 5.1.3, Addison-Wesley, Boston.
- Mann, H. B. (1945). On a test for randomness based on signs of differences, *Annals of Mathematical Statistics*, **16**, 193–199.
- Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to the Theory of Statistics*, 3rd ed., McGraw-Hill, Singapore.
- Stanley, R. P. (1997). *Enumerative Combinatorics*, Vol. 1, Cambridge University Press, New York.
- 土屋高宏, 中村永友(2007). ある種の並べ替え算法における分布について, 札幌学院大学商経論集, **24**(2), 31–47.

Eulerian Numbers in Modified Bucket Sorting and Its Related Distribution Theories

Takahiro Tsuchiya¹ and Nagatomo Nakamura²

¹Department of Mathematics, Josai University

²Department of Economics and General Education Center, Sapporo Gakuin University

This paper discusses the discrete probability distribution function induced from modified bucket sorting. The distribution is related to the unique sequence called “Eulerian Numbers” proposed by Euler. The sequence number obtained by modified bucket sorting is expressed as a recurrence formula. It is mathematically equivalent to the recurrence formula expression of Eulerian numbers. However, there are sensitive differences between a derivation process of our proposal and Eulerian numbers expressed as a number of permutation accents. The recurrence relation for the moments of the distribution is also given. The mean, the variance and higher order cumulants are derived using the relation. Interesting results of the relationship among the proposed, normal and uniform distributions are shown by asymptotic expansion. The efficiency of the proposed method is illustrated through numerical experiments.

Key words: Asymptotic expansion, cumulant, Eulerian distribution, normal distribution, uniform distribution.