

ブロードキャストコミュニケーションにおける Bajaj の一人勝ち残り公平なコイン投げ問題の平均回数 の漸近解

南山大学大学院数理情報研究科 須崎政文 (Masabumi Suzaki)

Graduate School of Mathematical Sciences and Information Engineering,
Nanzan University

南山大学数理情報学部 尾崎俊治 (Shunji Osaki)

Faculty of Mathematical Sciences and Information Engineering,
Nanzan University

1 はじめに

コイン投げは確率論で論じられる代表的モデルである。Bruss and O’Cinneide [1] によれば P.N. Bajaj は 1988 年の American Mathematical Society の集会で An Open Toss Problem というタイトルの次のようなコイン投げ問題を考えた：

N 人がコインを投げる。裏を出したものは除外され、表を出したものは再びコインを投げる。全員が裏あるいは表を出したら再び同じ人数でコインを投げる。コインは公平とする。このとき、最後の 1 人が残るまでの平均のコインを投げる回数はいくつになるだろう？

我々はこの問題を Bajaj の一人勝ち残り公平なコイン投げ問題と呼ぶ。この問題はブロードキャストコミュニケーションにおいて通信の際衝突が起きたときにも使われる。衝突を解消するためにコイン投げモデルが使われるのである。それは、“selection of a loser” [7] あるいは “leader election” [11, 3, 4, 5] と呼ばれている (Grabner and Prodingner [2])。

Bruss and O’Cinneide [1] はこの問題の平均回数を漸化式として求めた。その漸化式が複雑であり厳密解を求めることが困難であることから、母関数を用いて漸近解を求めることを試みたが、しかし漸近解を求めることはできなかった。

Prodingner [7] は P.N. Bajaj や Bruss and O’Cinneide [1] とは独立に同じ問題を研究した。ただし、実際はタイトル “How to select a loser” とあるように Prodingner は 1 人の勝者ではなく 1 人の敗者を選ぶ問題を研究した。しかし、コインの表裏が公平であることから 1 人の勝者を選ぶことと 1 人の敗者を選ぶことは全く同じ問題である。Prodingner [7] はこのコイン投げ問題の構造が Incomplete Trie (Knuth [8]) であると捉えた。また、このコイン投げ問題は Tree Protocol (Mathys and Flajolet [9]) でも使われていると指摘した。Prodingner [7] は漸化式の平均回数を母関数に加えて Rice’s Method と呼ばれる高度な解析手法を用いて漸近解を求めた。

伊藤ら [10] は前述の 2 論文とは独立に同じ問題を研究した。伊藤ら [10] は漸化式の平均回数を Bruss や Prodingner が用いた母関数ではなくアルゴリズム解析手法の一つであるブートストラッピング [11, 12] を用いて平均回数のオーダを求めた。また、伊藤ら [10] は Bajaj のコイン投げ問題の発展として、コイン投げを

じゃんけん置き換えたじゃんけん問題を考えている。伊藤ら [10] はこのじゃんけん問題のじゃんけんが終わるまでの平均回数のオーダーも求めている。

前述のとおり、Bajaj の一人勝ち残り公平なコイン投げ問題の平均回数の漸近解は母関数と Rice's Method を用いて Prodinger [7] によって求められている。しかし、Bruss and O'Conneide [1] が母関数を用いても正確な漸近解は求めることが出来なかったようにこの問題の漸化式は複雑な漸化式となっている。本論文は Bruss and O'Conneide [1] や Prodinger [7] が用いた母関数でも解くのが容易でないこの漸化式に、伊藤ら [10] が用いたブートストラッピングを適用する。伊藤ら [10] はオーダーを求めたが、我々はより精度の高い漸近解を求める。その漸近解は Prodinger の漸近解とは異なる式の漸近解であるが、数値計算によって N が十分に大きいと Prodinger の漸近解とも近づくことを示す。

2 Bajaj の一人勝ち残り公平なコイン投げ問題

Bajaj の一人勝ち残り公平なコイン投げ問題の定義は以下の通りである：

1. N 人から 1 人が残る。ここで $2 \leq N$ である。
2. コインの表裏を出す確率はそれぞれ $1/2$ とする。
3. 各コイン投げは独立である。
4. コインにより N 人から 1 人が残る回数を確率変数 C_N とする。
5. 1 回のコイン投げで残った人数を変数 m とする。ここで $1 \leq m \leq N-1$ である。
6. 参加者全員で一斉に同種のコインを投げる。
7. 裏を出した者は除外される。
8. 表を出した者は残って再びコインを投げる。
9. 参加者が全員表あるいは裏を出したら引き分けとして、まったく同じ参加者で再びコインを投げる。
10. $m = 1$ つまり残った人数が 1 人ならば目的が達成されるのでコイン投げは終了する。
11. $m > 1$ ならば m 人から 1 人が残るコイン投げ問題として続ける (再帰する)。

3 平均回数とその漸近解

コイン投げ平均回数は引き分けになる場合と勝者が決まる場合の排反事象から求めることができる。平均回数は漸化式として

$$E[C_N] = 1 + \frac{2}{2^N - 2} + \frac{1}{2^N - 2} \sum_{m=1}^{N-1} \binom{N}{m} E[C_m] \quad (1)$$

となる。ただし、 $E[C_0] = E[C_1] = 0$ である。この種の漸化式は Patricia tree recurrence と呼ばれる (Flajolet and Sedgewick [6])。

3.1 ブートストラッピングを用いた漸近解

ブートストラッピングとはアルゴリズム解析における複雑な漸化式の解析手法の一つであり、一般的には厳密解ではなく漸近解を求める解析手法である。また、ヒューリスティックな手法であり、それ故に様々な試行錯誤が必要になる。ブートストラッピングの解析手順は以下の通りである：

1. 漸化式をいくつかの数値で計算する.
2. 1 で得られた結果から上界と下界の不等式を推測する.
3. 2 で推測した上界と下界の不等式を漸化式に代入する.
4. 3 で得られた結果が (2) で推測した上界と下界の不等式よりも平均回数に近づいていることを証明する.
5. 3 と 4 を上界と下界が十分に漸近するまで繰り返し漸近解を得る.

ブートストラッピングの特徴は、一般的性質として粗い近似からでもブートストラッピングを繰り返すことによって徐々に近似の精度を上げられ、やがて漸近解を求められる点にある。ただし、上手くブートストラッピングの出発点を考えなければブートストラッピングを繰り返す途中で計算が難しくなってしまう可能性がある。

このブートストラッピングを式 (1) に適用する。最初に漸化式をコンピュータでいくつかの数値で計算すると、 $E[C_2] = 2.00$, $E[C_4] = 2.67$, $E[C_8] = 3.59$, $E[C_{16}] = 4.55$, $E[C_{32}] = 5.52$ となった。得られた結果から本論文では試行錯誤の結果、上界として $E[C_N] \leq N$, 下界として $E[C_N] > 0$ と推測する。この不等式の証明は下界は自明であり、上界は数学的帰納法を用いて容易に証明できる。次に推測した上界を漸化式に代入すると上界は

$$\begin{aligned}
 E[C_N] &= 1 + \frac{2}{2^N - 2} + \frac{1}{2^N - 2} \sum_{m=1}^{N-1} \binom{N}{m} E[C_m] \\
 &= 1 + \frac{2}{2^N - 2} + \frac{1}{2^N - 2} \sum_{m=2}^{N-1} \binom{N}{m} E[C_m] \\
 &\leq 1 + \frac{2}{2^N - 2} + \frac{1}{2^N - 2} \sum_{m=2}^{N-1} \binom{N}{m} m \\
 &= 1 + \frac{2}{2^N - 2} + N \frac{1}{2^N - 2} \sum_{m=1}^{N-2} \binom{N-1}{m} \\
 &= 1 + \frac{2}{2^N - 2} + \frac{N}{2} - \frac{N}{2^N - 2}
 \end{aligned} \tag{2}$$

となる。ここで、 $1 + \frac{2}{2^N - 2} + \frac{N}{2} - \frac{N}{2^N - 2} < N$ であるのは明らかである。

推測した下界を漸化式に代入すると下界は

$$\begin{aligned}
 E[C_N] &= 1 + \frac{2}{2^N - 2} + \frac{1}{2^N - 2} \sum_{m=1}^{N-1} \binom{N}{m} E[C_m] \\
 &= 1 + \frac{2}{2^N - 2} + \frac{1}{2^N - 2} \sum_{m=2}^{N-1} \binom{N}{m} E[C_m] \\
 &> 1 + \frac{2}{2^N - 2} + \frac{1}{2^N - 2} \sum_{m=2}^{N-1} \binom{N}{m} 0 \\
 &= 1 + \frac{2}{2^N - 2}
 \end{aligned} \tag{3}$$

となる。ここで、 $0 < 1 + \frac{2}{2^N - 2}$ であるのは明らかである。

以上のように、推測した上界と下界の不等式をそれぞれ漸化式に代入すると、新しく求められた不等式が上界も下界もより平均回数に近い不等式となる。

ここで、一般回数 l 回のブートストラッピングをおこなうと上界は

$$\begin{aligned}
 E[C_N] &< l + \frac{N}{2^l} \\
 &- Nl \sum_{i=1}^{l-1} \sum_{j=0}^{2^{i-1}-1} \frac{1}{2^i} \left[\left(1 - \frac{2j+1}{2^i}\right)^{N-1} - \left(1 - \frac{2j+2}{2^i}\right)^{N-1} \right] \\
 &+ N \sum_{i=1}^{l-1} \sum_{j=0}^{2^{i-1}-1} i \frac{1}{2^i} \left[\left(1 - \frac{2j+1}{2^i}\right)^{N-1} - \left(1 - \frac{2j+2}{2^i}\right)^{N-1} \right]
 \end{aligned} \tag{4}$$

となる。下界は

$$\begin{aligned}
 E[FC_N] &> l - Nl \sum_{i=1}^{l-1} \sum_{j=0}^{2^{i-1}-1} \frac{1}{2^i} \left[\left(1 - \frac{2j+1}{2^i}\right)^{N-1} - \left(1 - \frac{2j+2}{2^i}\right)^{N-1} \right] \\
 &+ N \sum_{i=1}^{l-1} \sum_{j=0}^{2^{i-1}-1} i \frac{1}{2^i} \left[\left(1 - \frac{2j+1}{2^i}\right)^{N-1} - \left(1 - \frac{2j+2}{2^i}\right)^{N-1} \right]
 \end{aligned} \tag{5}$$

となる。

ここで、上界と下界の差は $\frac{N}{2^l}$ であるのでこの値が十分に小さくなればなるほど、上界と下界は漸近して、平均回数の漸近解が求まる。すなわち平均回数は

$$\begin{aligned}
 E[C_N] &\sim l - Nl \sum_{i=1}^{l-1} \sum_{j=0}^{2^{i-1}-1} \frac{1}{2^i} \left[\left(1 - \frac{2j+1}{2^i}\right)^{N-1} - \left(1 - \frac{2j+2}{2^i}\right)^{N-1} \right] \\
 &+ N \sum_{i=1}^{l-1} \sum_{j=0}^{2^{i-1}-1} i \frac{1}{2^i} \left[\left(1 - \frac{2j+1}{2^i}\right)^{N-1} - \left(1 - \frac{2j+2}{2^i}\right)^{N-1} \right]
 \end{aligned} \tag{6}$$

となる。

そこで、本論文ではブートストラッピングをおこなう回数を $l = 2[\log_2 N] + 21 = 2\log_2 N - 2\{\log_2 N\} + 21$ とする。すると、上界と下界の差 $\frac{N}{2^l}$ は

$$\begin{aligned}
 \frac{N}{2^l} &= \frac{N}{2^{2\log_2 N - 2\{\log_2 N\} + 21}} \\
 &= \frac{1}{2^{21}} \frac{N}{2^{2\log_2 N - 2\{\log_2 N\}}} \\
 &= \frac{N 2^{2\{\log_2 N\}}}{2^{21} N^2} \\
 &= \frac{2^{2\{\log_2 N\}}}{2^{21} N}
 \end{aligned} \tag{7}$$

となる。よって N が十分に大きくなれば上界と下界は漸近して、平均回数の漸近解が求まる。

ここで、式 (6) の

$$Nl \sum_{i=1}^{l-1} \sum_{j=0}^{2^{i-1}-1} \frac{1}{2^i} \left[\left(1 - \frac{2j+1}{2^i}\right)^{N-1} - \left(1 - \frac{2j+2}{2^i}\right)^{N-1} \right] \tag{8}$$

を $i = \lfloor \log_2 N \rfloor = \log_2 N - \{\log_2 N\}$ を中心とした式に変換すると

$$\begin{aligned} & Nl \sum_{i=1}^{l-1} \sum_{j=0}^{2^{i-1}-1} \frac{1}{2^i} \left[\left(1 - \frac{2j+1}{2^i}\right)^{N-1} - \left(1 - \frac{2j+2}{2^i}\right)^{N-1} \right] \\ & \sim l \sum_{i=-\lfloor \log_2 N \rfloor+1}^{\lfloor \log_2 N \rfloor+20} \frac{\left(\frac{1}{2}\right)^i 2^{\{\log_2 N\}}}{e^{\left(\frac{1}{2}\right)^i 2^{\{\log_2 N\}} + 1}} \\ & \approx l \int_{-\lfloor \log_2 N \rfloor+1}^{\lfloor \log_2 N \rfloor+20} \frac{\left(\frac{1}{2}\right)^i 2^{\{\log_2 N\}}}{e^{\left(\frac{1}{2}\right)^i 2^{\{\log_2 N\}} + 1}} di \\ & = -l \log_2 e \int_{\frac{N}{2^{\{\log_2 N\}_2}}^{\frac{2^{\{\log_2 N\}}}{2^{20N}}} \frac{2^{\{\log_2 N\}}}{e^{2^{\{\log_2 N\}}t} + 1}} dt, \end{aligned}$$

ここで積分の公式を用いて

$$\begin{aligned} & = -l \log_2 e \int_{\frac{N}{2^{\{\log_2 N\}_2}}^{\frac{2^{\{\log_2 N\}}}{2^{20N}}} \frac{2^{\{\log_2 N\}}}{e^{2^{\{\log_2 N\}}t} + 1}} dt \\ & = -l \log_2 e \left[\log_e \frac{e^{2^{\{\log_2 N\}}t}}{e^{2^{\{\log_2 N\}}t} + 1} \right]_{\frac{N}{2^{\{\log_2 N\}_2}}^{\frac{2^{\{\log_2 N\}}}{2^{20N}}} \\ & = -l \cdot \log_2 e \log_e \frac{e^{\frac{(2^{\{\log_2 N\}})^2}{2^{20N}}}}{e^{\frac{(2^{\{\log_2 N\}})^2}{2^{20N}} + 1}} + l \cdot \log_2 e \log_e \frac{e^{\frac{N}{2}}}{e^{\frac{N}{2}} + 1} \\ & \sim -l \cdot \log_2 e \log_e \frac{1}{2} + l \cdot \log_2 e \log_e 1 \\ & = l \cdot \log_2 e \cdot \log_e 2 \\ & = l \end{aligned} \tag{9}$$

となる.

次に式(6)の

$$N \sum_{i=1}^{l-1} \sum_{j=0}^{2^{i-1}-1} i \frac{1}{2^i} \left[\left(1 - \frac{2j+1}{2^i}\right)^{N-1} - \left(1 - \frac{2j+2}{2^i}\right)^{N-1} \right] \tag{10}$$

を $i = \lfloor \log_2 N \rfloor = \log_2 N - \{\log_2 N\}$ を中心とした式に変換すると

$$\begin{aligned} & N \sum_{i=1}^{l-1} \sum_{j=0}^{2^{i-1}-1} i \frac{1}{2^i} \left[\left(1 - \frac{2j+1}{2^i}\right)^{N-1} - \left(1 - \frac{2j+2}{2^i}\right)^{N-1} \right] \\ & \sim \lfloor \log_2 N \rfloor \sum_{i=-\lfloor \log_2 N \rfloor+1}^{\lfloor \log_2 N \rfloor+20} \frac{\left(\frac{1}{2}\right)^i 2^{\{\log_2 N\}}}{e^{\left(\frac{1}{2}\right)^i 2^{\{\log_2 N\}} + 1}} + \sum_{i=-\lfloor \log_2 N \rfloor+1}^{\lfloor \log_2 N \rfloor+20} \frac{i \left(\frac{1}{2}\right)^i \cdot 2^{\{\log_2 N\}}}{e^{\left(\frac{1}{2}\right)^i 2^{\{\log_2 N\}} + 1}} \\ & \sim \lfloor \log_2 N \rfloor + \sum_{i=-\lfloor \log_2 N \rfloor+1}^{\lfloor \log_2 N \rfloor+20} \frac{i \left(\frac{1}{2}\right)^i \cdot 2^{\{\log_2 N\}}}{e^{\left(\frac{1}{2}\right)^i 2^{\{\log_2 N\}} + 1}} \end{aligned} \tag{11}$$

となる.

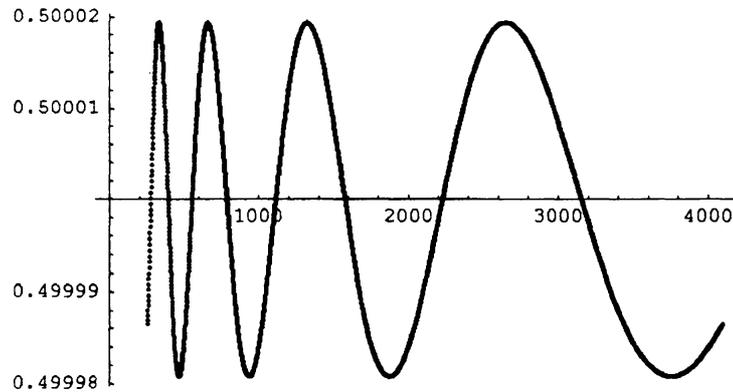


図1 式(14)を $N = 2^8, \dots, 2^{12}$ でプロット

式(6)に式(9)と式(11)を代入すると漸近解

$$E[C_N] \sim \lfloor \log_2 N \rfloor + \sum_{i=-\lfloor \log_2 N \rfloor + 1}^{\lfloor \log_2 N \rfloor + 20} \frac{i \left(\frac{1}{2}\right)^i \cdot 2^{\lfloor \log_2 N \rfloor}}{e^{\left(\frac{1}{2}\right)^i \cdot 2^{\lfloor \log_2 N \rfloor}} + 1} \quad (12)$$

が求まる.

3.2 Prodingerの漸近解との比較

Prodingerの漸近解は

$$\log_2 N + 0.5 - \frac{1}{L} \sum_{k \neq 0} \zeta \left(1 - \frac{2k\pi i}{L}\right) \Gamma \left(1 - \frac{2k\pi i}{L}\right) e^{2k\pi i \log_2 N} \quad (13)$$

である. ここで, $L = \log 2$, $\zeta(\cdot)$ は Riemann の ζ 関数, $\Gamma(\cdot)$ は Γ 関数, i は虚数単位の i である.

Prodingerの漸近解(13)と本論文の漸近解(12)を比較するため, 漸近解(13)から $\log_2 N$ を引いた

$$0.5 - \frac{1}{L} \sum_{k \neq 0} \zeta \left(1 - \frac{2k\pi i}{L}\right) \Gamma \left(1 - \frac{2k\pi i}{L}\right) e^{2k\pi i \log_2 N} \quad (14)$$

に注目する. 図1は式(14)を $N = 2^8, \dots, 2^{12}$ でプロットしたものである.

次に, 本論文の漸近解(12)から $\log_2 N$ を引いた

$$\begin{aligned} & - \{\log_2 N\} \\ & + \sum_{i=-\lfloor \log_2 N \rfloor + 1}^{\lfloor \log_2 N \rfloor + 20} \frac{i \left(\frac{1}{2}\right)^i \cdot 2^{\lfloor \log_2 N \rfloor}}{e^{\left(\frac{1}{2}\right)^i \cdot 2^{\lfloor \log_2 N \rfloor}} + 1} \end{aligned} \quad (15)$$

に注目する. 図2は式(15)を $N = 2^8, \dots, 2^{12}$ でプロットしたものである.

図1と図2は一見すると全く同じ図に見えるが, 実際にどの程度の誤差があるか検証する. 図3はProdinger

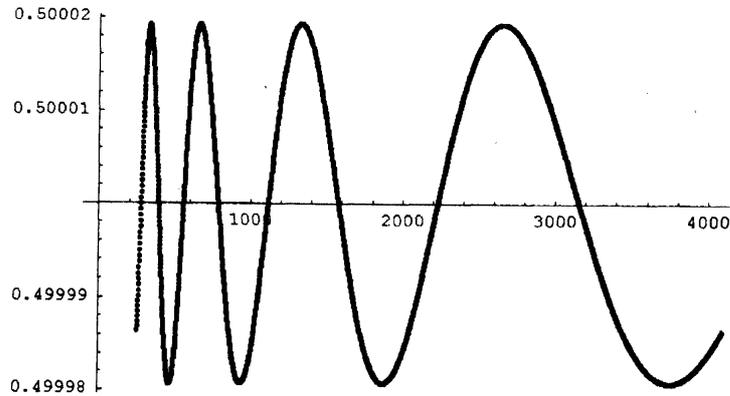


図2 式(15)を $N = 2^8, \dots, 2^{12}$ でプロット

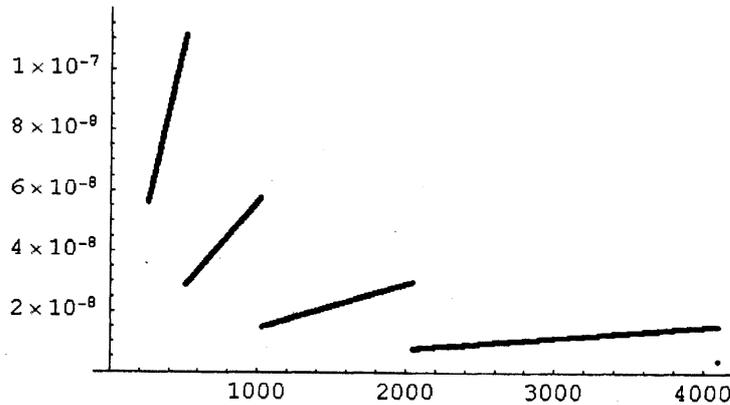


図3 式(16)を $N = 2^8, \dots, 2^{12}$ でプロット

の漸近解 (13) から本論文の漸近解 (12) を引いた

$$\begin{aligned}
 & 0.5 \\
 & - \frac{1}{L} \sum_{k \neq 0} \zeta \left(1 - \frac{2k\pi i}{L} \right) \Gamma \left(1 - \frac{2k\pi i}{L} \right) e^{2k\pi i \log_2 N} \\
 & + \{\log_2 N\} - \sum_{i=-\lfloor \log_2 N \rfloor + 1}^{\lfloor \log_2 N \rfloor + 20} \frac{i \left(\frac{1}{2}\right)^i \cdot 2^{\{\log_2 N\}}}{e^{\left(\frac{1}{2}\right)^i \cdot 2^{\{\log_2 N\}} + 1}}
 \end{aligned} \tag{16}$$

を $N = 2^8, \dots, 2^{12}$ でプロットしたものである。図より誤差が $\log_2 N$ の周期で減少している。よって、本論文の漸近解は式は異なるが Prodinger の漸近解と N が十分に大きいと近づくのである。

4 おわりに

本論文はブートストラッピングを用いて Bajaj の一人勝ち残り公平なコイン投げ問題の平均回数の漸近解を求めた。その結果、Prodinger の漸近解とは異なる式の漸近解が求まった。ただし、図3より数値計算に

よって本論文の漸近解 (12) は Prodinger の漸近解 (13) と N が十分に大きいと近づくことが分かった。

また、本論文の漸近解は式に $\lfloor \log_2 N \rfloor$ という不連続な関数を含んでいるにもかかわらず滑らかな関数となる点を強調したい。すなわち、Prodinger の漸近解とは明らかに式の構成が異なっており、別の表現になっているのである。

この種の漸化式、すなわち Patricia tree recurrence の解析は Mellin transform [6] が一般的であると言われている (Sedgewick and Flajolet [12])。本論文は Patricia tree recurrence の解析をブートストラッピングでおこなったものであり、本論文の結果よりブートストラッピングが将来のアルゴリズム解析手法の多様性の一端になることを期待したい。

謝辞 本研究の一部は文部科学省科学研究費補助金基盤研究 (B)20330068, (C)18510138 および 2009 年度南山大学パツへ研究奨励金 I-A-2 による助成のもとで行われたものである。

参考文献

- [1] F.T. Bruss and C.A. O’Cinneide, *On the maximum and its uniqueness for geometric random samples*, Journal of Applied Probability, Vol. 27, No. 3, pp.598-610, 1990.
- [2] P.J. Grabner and H. Prodinger, *Sorting algorithms for broadcast communications: Mathematical analysis*, Theoret. Comput. Sci., Vol. 289, No. 1, pp.51-67, 2002.
- [3] C. Lavault and G. Louchard *Asymptotic analysis of a leader election algorithm*, Theoret. Comput. Sci., Vol. 359, Nos. 1-3, pp.239-254, 2006.
- [4] J.A. Fill, H.M. Mahmoud and W. Szpankowski, *On the distribution for the duration of a randomized leader election algorithm*, Ann. Appl. Probab., Vol. 6, No. 4, pp.1260-1283, 1996.
- [5] S. Janson and W. Szpankowski, *Analysis of an asymmetric leader election algorithm* The Electronic journal of combinatorics 4, #R17, 1997.
- [6] P. Flajolet and R. Sedgewick, *Mellin transforms and asymptotics: Finite differences and rice’s integrals*, Theoret. Comput. Sci., Vol. 144, Nos. 1-2, pp.101-124, 1995.
- [7] H. Prodinger, *How to select a loser*, Discrete Mathematics, Vol. 120, Nos. 1-3, pp.149-159, 1993.
- [8] D.E. Knuth, *The Art of Computer Programming*, Vol. 3, 2nd ed., Addison-Wesley, Reading, Mass, 1998.
- [9] P. Mathys and P. Flajolet, *Q-ary collision resolution algorithms in random-access systems with free or blocked channel access*, IEEE Trans. Inform., Theory 31, pp.217-243, 1985.
- [10] 伊藤暁, 井上克司, 王躍, 岡崎世雄, ジャンケンの計算量, 電子情報通信学会論文誌 (D), Vol. J86, No. 7, pp.452-457, 2003.
- [11] R. Sedgewick and P. Flajolet, *An Introduction to the Analysis of Algorithms*, Addison-Wesley, Reading, Mass., 1996.
- [12] R. Graham, D.E. Knuth and O. Patashink, *Concrete Mathematics: A Foundation for Computer Science*, 2nd ed., Addison-Wesley, 1994.