有限回のコイン投げによる連続一様分布の最適な近似

The Optimal Approximation of the Continuous Uniform Distribution By Finite Number of Coin Tossing

（平成18年9月受理）

大槻 正伸*（OHTSUKI Masanobu）

Abstract

Random numbers with the uniform distribution on [0,1] is often used in computer simulations such as the Monte Carlo Method. And in computers, such pseudo random numbers are made by a 0,1bits sequence. In this paper, we discuss about the optimal method of generating the pseudo random numbers from a result of finite number of honest coin tossing——i.e. random 0,1 bits sequence.

We introduce at first two distances $d_1(D_1, D_2)$, $d_2(D_1, D_2)$ in natural way, where D_1, D_2 are two (discrete or continuous) distributions on [0,1], and then find the optimal method for generating random numbers in terms of d_1 and d_2 from honest coin tossing.

Finally we can show that the optimal method in terms of d_1 is also optimal in terms of d_2.

1．はじめに

本論文では「理想的なコインを何回か（有限回）投げるという試行から、[0,1] 上の一様分布になるべくよくしたがう(擬似)乱数を発生させるにはどうしたらよいか」という問題について考察する。すなわち、「コイン投げ（i.e.ランダムなビット列）から擬似乱数をつくる最良の方法」について考察する。

以下でこの問題が出てきた背景について述べる。

（背景1）

[0,1]上の一様分布（以下、「一様分布」という場合は[0,1]上の一様分布を意味するものとする）はモンテカルロ法などの様々なコンピュータシミュレーションに用いられる基本的な確率分布である。一様分布以外の確率分布を用いる場合も、基本的には一様分布を用いてつくることが多い[4][5][7]。その意味で一様分布はシミュレーションを行うのに基本的な確率分布である。

しかし、シミュレーションで用いる一様分布にしたがう乱数をコンピュータで発生させることは簡単ではなく、よい一様分布にしたがう乱数の発生はシミュレーションを行うには重要な技術となる。

コンピュータ内で一様分布を発生させる際には、何らかの離散システム（例えばあるメモリ内の0，1のビット列等）をもとに（擬似）乱数つくることが考えられる。その際には、例えばメモリ内のビット列が「01101100」であれば単純に $(0.01101100)_2$（整数部「0.」を頭につけ、2進数解釈する）＝ $\frac{1}{2^2}+\frac{1}{2^3}+\frac{1}{2^5}+\frac{1}{2^6}=0.421875$ を乱数候補とする、というようにすることが自然であるが、これで確かに、ランダムな8ビット列の情報から得られる一番よい一様乱数近似になっているか、というような議論はほとんどなされていない。

（本論文では、この方法はある意味で最適ではないことが示される（定理1））

また、従来の話題の多くは「どのようにして、決定論的な過程から一様乱数に近い数列を発生できるか」に重点がおかれている。例えば、整数 a,b,m,X_0 を適当に定めて、$X_{n+1}=aX_n+b \pmod m$、により（0～(m-1)の範囲の）整数の擬似乱数を発生させ $\frac{X_n}{m}$ (n=1,2,3,…) を（擬似）一様乱数とする（Lehmer法）、等がある[4][5]。これも、$\frac{0}{m}\sim\frac{m-1}{m}$ がランダムにでたとき、それらを自然にそのまま一様乱数と見るのであって、m個の数の集合 $\{\frac{0}{m},\frac{1}{m},\frac{2}{m},\cdots,\frac{m-1}{m}\}$ の値をとる離散確率変数 X（ただし P[X=k/m]=1/m, k=0,1,..,m-1）があったら、その後どのようにするのが本当に正確な一様乱数を得る方法か、等についての議論ではない。

*福島工業高等専門学校　電気工学科（いわき市平上荒川字長尾30）

正確なランダムビット列発生後の一様乱数生成について数学的に抽象化すると、当初示した問題がでてくる。

（背景２）

まず、正方形と円により円周率πの値を推定するモンテカルロ法を考える[(2)(5)(6)(8)]。これは一様分布にしたがう 2 つの乱数 x, y をとり、点(x, y)を正方形$[0,1]^2$にランダムにばらまいた点と見る。このようにしてｎ個の点を正方形内にランダムにばらまき、そのうち内接円内に入った点を数えそれがｍ個だったとき、$\pi \fallingdotseq \frac{4m}{n}$で推定するものである。

ここで、このモンテカルロ法は「乱数（一様分布にしたがう乱数 2n 個）の情報がπの値に関する近似情報に変換された」と見ることができる。

さて、それでは、この場合どれだけの情報量がどれだけの情報量に変換されたのであろうか、というような議論をしようとすると次のような壁につき当たってしまう。

もともと、真に正確な一様乱数１個は、無限の 0, 1 列の集合 $\{0,1\}^{\infty}$の要素で、０も１も確率 $\frac{1}{2}$ で出現しているものと考えることができるから、∞ビットの情報をもっている。

実際、例えば、正確な乱数 1 個を 2 進数表現し、次の無限長ビット列が得られたとする。

0.10101101001010100101010100011110101･･･

これを、例えば 10 ビットずつ区切って、

0.1010110100 | 1010100101 | 0101000111 | 10101･･･

として、$x_1=(0.1010110100)_2$, $x_2=(0.1010100101)_2$, $x_3=(0.0101000111)_2$, … というようにすると、いくつもの（一様分布にしたがう）擬似乱数が得られる（区切り幅を 10 ビットでなく 20 ビット、あるいはそれ以上に増やせばより精度のよい擬似乱数が得られる)。

これらを用いれば、「一様分布にしたがう乱数１個」から、無限個の非常によい擬似乱数をつくることができ、したがって、モンテカルロ法によりいくらでもよいπの近似値が得られることになる。

したがって理論的には、「一様乱数にしたがう確率変数の実現値１個の情報量は無限大であるから、乱数 1 個で、モンテカルロ法によりいくらでもよいπの近似情報を得ることができる」ということになる。

正確な一様乱数とは、このように非常に多くの情報量を持つものであり、人間の力では正確なものは発生できないものと考えられる。

つまり[0, 1]上の一様乱数を情報量の基本にとっては「乱数情報がπの値に関する近似情報に変換された」というような見方での様々な解析はできなくなる。

そこで、「理想的なコイン投げ」を基本的な情報とみて、「コイン投げ（ｎ回)」→「0, 1 のｎビット列」→「一様乱数の近似」→モンテカルロ法によるπの近似、というように考えれば、もともとの情報量もShannon の情報理論[(1)(3)]で計量することができ、上記のような情報量的解析も可能となると考えられる。

それでは、有限回のコイン投げからどのように一様乱数を生成するのが最もよい方法なのか、という問題が出てくる。これは、離散分布と連続な一様分布との関係を明らかにする問題である。

以上のように当初に示した問題は上記 2 つの背景をもった問題であると見ることができる。

特に背景 1 では、標本空間の大きさが大きくなれば、一様乱数生成が最適かどうかはあまり問題にならなくなるが、背景 2 の情報理論的考察をする際に以下の議論が解析の基礎になると考えられる。

本論文の目的は、以上の問題に対してある程度の解答を与えることである。

それには、まず「ある離散分布が一様分布に近い」という概念を厳密に定義し、その上で、ｎ回のコイン投げからなるべく「一様乱数に近い」分布を得る、というような議論が可能になる。

以下 2. では、2 つの確率分布に対して近さの基準、すなわち 2 つの確率分布 D_1, D_2の「距離」2 種類 $d_1(D_1,D_2)$と $d_1(D_1,D_2)$を定義する。これは、離散的な分布と連続的な分布の間でもその距離が計量できるものでなくてはならない。

3. では、確率空間の要素数を一定ｋにした場合、どのような離散分布が、距離 d_1の意味で一様乱数に最も近くなるか、について結論を導く。

4. では、同様に距離 d_2の意味で、どのような離散分布が一様乱数に最も近くなるか、について調べる。

また、この結果から「コイン投げから距離　d_1の意味で最適な離散分布を得る方法は距離 d_2の意味でも最適な離散分布を得る方法になる」ことが示される。

２．２つの分布の距離

以下$P_D[A]$で事象Ａの分布Ｄにおける確率を表す。また、離散分布を $\begin{pmatrix} x_1 & x_2 & \cdots & x_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}$ で表す。

これは Fig. 1 のように、$P_D[X=x_i]=p_i$

$(i=1,2,..,k)$という離散確率分布を意味する。

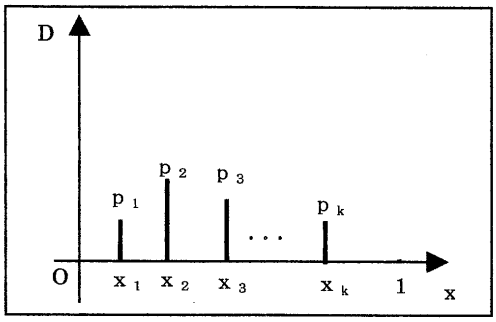

Fig.1 An example of a discrete distribution D

当然ながら、$\Sigma p_i = 1$ となる。

今回は、[0,1]上の一様分布との比較をすることが目的であるから、$0 \leqq x_1 < x_2 < \ldots < x_k \leqq 1$ を仮定しておく。

D_1 と D_2 を 2 つの、区間 [0,1] 上の確率分布とする。このとき 2 つの分布に対して、距離 $d_1(D_1, D_2)$、および、距離 $d_2(D_1, D_2)$ を次で定義する。

これらが数学的な距離になることは後に示される。

【定義 1】(d_1, d_2)

(1) $F_1(x), F_2(x)$ をそれぞれ D_1, D_2 の分布関数とする。

$$d_1(D_1, D_2) \overset{\text{def}}{=} \int_0^1 | F_1(x) - F_2(x) | \, dx$$

(2) $$d_2(D_1, D_2) = \sup_{[a,b] \subset [0,1]} \{ d_{[a,b]}(D_1, D_2) \}$$

ただし、$$d_{[a,b]}(D_1, D_2) \overset{\text{def}}{=} | P_{D1}[[a,b]] - P_{D2}[[a,b]] |$$

□(定義 1)

以下で、U は ([0,1]上の) 一様分布を意味するものとする。

＜例 1＞　この例では $D_{ex} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$　とする。

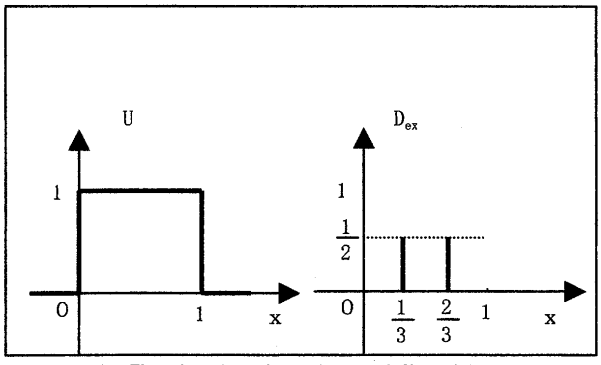

Fig.2 The density functions of U and D_{ex}

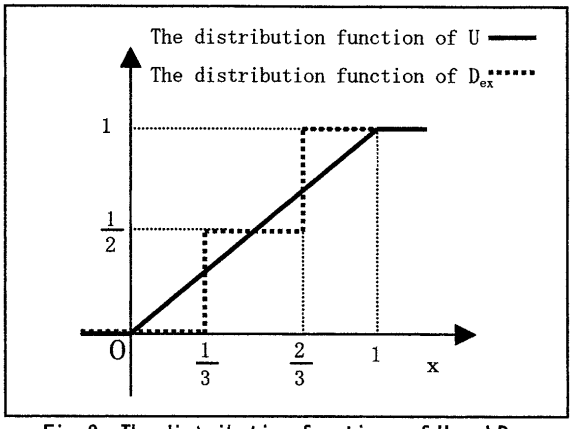

Fig.3 The distribution functions of U and D_{ex}

$d_1(U, D_{ex})$ ＝Fig3. の 4 つの三角形の面積の和 $= \frac{5}{36}$

$d_2(U, D_{ex}) = \max\{ \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{3}, \frac{1}{2}, \frac{1}{3} \} = \frac{2}{3}$

$d_2(U, D_{ex}) = \frac{2}{3}$　であることは次のように示せる。

$A = \{[a,b] ; [a,b] \subset [0,1]\}$ を次の 6 パターンに分類する。

$A_1 = \{[a,b] ; 0 \leqq a < b < \frac{1}{3}\}$

$A_2 = \{[a,b] ; 0 \leqq a \leqq \frac{1}{3} < b < \frac{2}{3}\}$

$A_3 = \{[a,b] ; 0 \leqq a \leqq \frac{1}{3}, \frac{2}{3} \leqq b \leqq 1\}$

$A_4 = \{[a,b] ; \frac{1}{3} < a < b < \frac{2}{3}\}$

$A_5 = \{[a,b] ; \frac{1}{3} < a \leqq \frac{2}{3} < b \leqq 1\}$

$A_6 = \{[a,b] ; \frac{2}{3} < a < b \leqq 1\}$

$$\sup_{[a,b] \in A_1} \{ d_{[a,b]}(U, D) \} = \frac{1}{3}, \quad \sup_{[a,b] \in A_2} \{ d_{[a,b]}(U, D) \} = \frac{1}{2}$$

等々、上記 6 パターンについて考えればよい。

□ (例 1)

直感的には、d_1 は D_1, D_2 の分布関数の一致しない部分の面積でその距離を計量しようというものであり、d_2 は区間 [a,b] を動かし、2 つの分布でそこに入る確率の差の最も大きくなるもので計量しようとするものである。d_1, d_2 は両方とも、離散分布、連続分布に限らず 2 つの分布の距離を計量することができるのは明らかである。

【性質 1】d_1, d_2 は数学的距離である。

＜証明＞

(1) (*0) $d_1(D_1, D_2) \geqq 0$、(*1) $d_1(D, D) = 0$

(*2) $d_1(D_1, D_2) = d_1(D_2, D_1)$ は明らか。

(*3) $d_1(D_1, D_2) + d_1(D_2, D_3) \geqq d_1(D_1, D_3)$

$\because D_1, D_2, D_3$ の分布関数をそれぞれ、$F_1(x)$,

$F_2(x), F_3(x)$とすると、
各$x \in [0,1]$において、
$|F_1(x)-F_2(x)|+|F_2(x)-F_3(x)| \geqq |F_1(x)-F_3(x)|$となっていること、および【定義１】(1)よりすぐにわかる。

(2) (*0) $d_2(D_1,D_2) \geqq 0$、(*1) $d_2(D,D)=0$
(*2) $d_2(D_1,D_2)=d_2(D_2,D_1)$は明らか。
(*3) $d_2(D_1,D_2)+d_2(D_2,D_3) \geqq d_2(D_1,D_3)$
∵任意の区間[a,b]において、

$$d_{[a,b]}(D_1,D_2)+d_{[a,b]}(D_2,D_3)=$$
$$|P_{D1}[[a,b]]-P_{D2}[[a,b]]|+$$
$$|P_{D2}[[a,b]]-P_{D3}[[a,b]]| \geqq$$
$$|P_{D1}[[a,b]]-P_{D3}[[a,b]]|=d_{[a,b]}(D_1,D_3)$$

より、

$$d_2(D_1,D_3)=\sup_{[a,b]\subset[0,1]}\{d_{[a,b]}(D_1,D_3)\} \leqq$$

$$\sup_{[a,b]\subset[0,1]}\{d_{[a,b]}(D_1,D_2)+d_{[a,b]}(D_2,D_3)\}$$

$$\leqq \sup_{[a,b]\subset[0,1]}\{d_{[a,b]}(D_1,D_2)\}+\sup_{[a,b]\subset[0,1]}\{d_{[a,b]}(D_2,D_3)\}$$

$$=d_2(D_1,D_2)+d_2(D_2,D_3)$$ □（性質１）

３．距離d_1の意味での最適な一様分布近似

$D=\begin{pmatrix} x_1 & x_2 & \cdots & x_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}$, U：[0,1] 上の一様分布とする（以下ではkを固定して考える）。
ただし、D において、$p_i \neq 0$ (i.e. $p_i>0$) としておく。

【補題A】$x_1\ x_2\ \cdots\ x_k$を一般にとって固定する。$d_1(D,U)$は次の$p_1, p_2, \cdots, p_k$のときに最小になる。

$$p_1=\frac{x_1+x_2}{2} \qquad p_2=\frac{x_2+x_3}{2}-p_1$$

$$p_i=\frac{x_i+x_{i+1}}{2}-(p_1+\cdots+p_{i-1})$$
$$(i=2,3,\cdots,k-1)$$

＜証明＞
まず、$D_1=\begin{pmatrix} x_1 & x_2 & \cdots & x_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}$

ただし、$p_1=x_1$, $p_2+p_1=x_2$ (i.e. $p_2=x_2-p_1=x_2-x_1$), $p_1+p_2+p_3=x_3\cdots$ という分布を考える（Fig.4）。

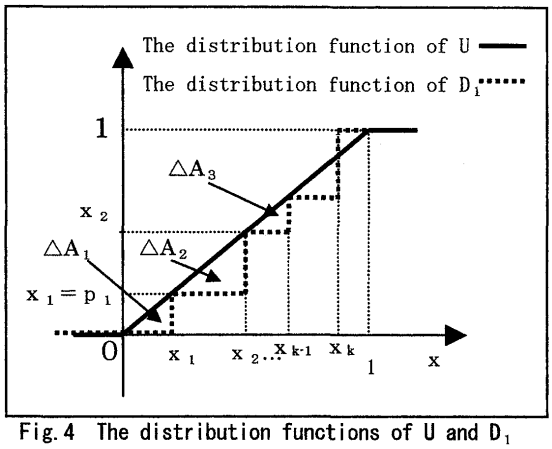

Fig.4 The distribution functions of U and D_1

このD_1を基準に考えると、$d_1(D,U)$を最小にするp_1は、$x_1 \leqq p_1 \leqq x_2$ でなくてはならないことがわかる。なぜなら$d_1(D_1,U)$はFig.4の三角形A_1, A_2, A_2, ... の面積の和であるが、$p_1<x_1$であればA_2が三角形でなく台形となり、その面積は三角形A_2の面積より大きくなるからこの場合$d_1(D_1,U)$を最小にし得ない。

また、$p_1>x_2$の場合も、三角形A_2に相当する部分が、三角形A_2の面積よりも面積の大きい台形（ただしUの分布関数の上側）となる。
以下、同様に考えて、
$d_1(D,U)$を最小にするには、まずは、

$$x_1 \leqq p_1 \leqq x_2$$
$$x_2 \leqq p_1+p_2 \leqq x_3$$
$$\vdots$$
$$x_i \leqq p_1+p_2+\cdots+p_i \leqq x_{i+1}$$
$$\vdots$$
$$x_{k-1} \leqq p_1+p_2+\cdots+p_{k-1} \leqq x_k$$

が成り立っていなくてはならないことがわかる

さて、$x_1 \leqq p_1 \leqq x_2$のとき、

$$\int_{x_1}^{x_2}|F_U(x)-F_D(x)|\,dx=\frac{(p_1-x_1)^2}{2}+\frac{(x_2-p_1)^2}{2}$$

($F_U(x), F_D(x)$はそれぞれ、U、Dの分布関数とする。)

であるから、これを最小にするp_1は、$p_1=\frac{x_1+x_2}{2}$
となる。

そして、そのとき$\int_{x_1}^{x_2}|F_U(x)-F_D(x)|\,dx=\frac{(x_2-x_1)^2}{4}$
となる。　以下、全く同様に考えて、$d_1(D,U)$を最小にする$p_2, \cdots, p_k$は、

$p_1+p_2=\frac{x_2+x_3}{2}$　より、$p_2=\frac{x_2+x_3}{2}-p_1$
$\vdots$

$p_i=\frac{x_i+x_{i+1}}{2}-(p_1+\cdots+p_{i-1})$

$(i=2,3,\cdots,k-1)$ となる。

そして、このとき、

$$\int_0^1 |F_U(x)-F_D(x)|\,dx=\frac{x_1^2}{2}+\frac{(1-x_k)^2}{2}+\sum_{i=2}^{k}\frac{(x_i-x_{i-1})^2}{4}$$

□（補題A）

【定理1】$d_1(D,U)$を最小にする分布

$D_{1min}=\begin{pmatrix} x_1 & x_2 & \cdots & x_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}$ は、

$$x_i=\frac{2i-1}{2k},\quad p_i=\frac{1}{k}\qquad (i=1,2,\cdots,k)$$

という分布である。

また、このときの $d_1(D_{1min},U)=\frac{1}{4k}$ となる。

＜証明＞

$$\phi(x_1,x_2,\cdots,x_k)\overset{def}{=}\frac{x_1^2}{2}+\frac{(1-x_k)^2}{2}+\sum_{i=2}^{k}\frac{(x_i-x_{i-1})^2}{4}$$

として、ϕを最小にする $x_1,x_2,\cdots,x_k$ を求める。ϕは$x_1,x_2,\cdots,x_k$に関する2次式であるから、

$\frac{\partial\phi}{\partial x_i}=0$ $(i=1,2,\cdots,k)$ の連立方程式を解く

ことにより、ϕを最小にする$x_1,x_2,\cdots,x_k$が求まる。

すなわち、

$$\begin{cases} \frac{\partial\phi}{\partial x_1}=x_1-\frac{1}{2}(x_2-x_1)=0 \\ \frac{\partial\phi}{\partial x_i}=\frac{1}{2}(2x_i-x_{i-1}-x_{i+1})=0 \quad (i=2,\cdots,k-1) \\ \frac{\partial\phi}{\partial x_k}=\frac{1}{2}(x_k-x_{k-1})-(1-x_k)=0 \end{cases}$$

これは、

$$\begin{cases} 3x_1-x_2=0 & (1) \\ -x_1+2x_2-x_3=0 & (2) \\ -x_2+2x_3-x_4=0 & (3) \\ -x_3+2x_4-x_5=0 & (4) \\ \vdots & \\ -x_{k-2}+2x_{k-1}-x_k=0 & (k-1) \\ -x_{k-1}+3x_k=2 & (k) \end{cases}$$

となる。

$$\begin{cases} 3x_1-x_2=0 & (1') \\ \quad これを x_2=3x_1 として(2)に代入すると、 & \\ 5x_1-x_3=0 & (2') \\ \quad これを、x_3=5x_1 として(3)に代入すると、 & \\ 7x_1-x_4=0 & (3') \\ \quad 以下同様にして（正確には数学的帰納法により）、 & \\ (2i-1)\,x_1=x_i & ((i-1)') \quad (i=2,3,..,k) \\ -(2k-3)\,x_1+3(2k-1)x_1=2 & (k') \end{cases}$$

(k')より、$4kx_1=2$ i.e. $x_1=\frac{1}{2k}$

したがって、 $x_i=\frac{2i-1}{2k}$ $(i=1,2,\cdots k)$

補題Aより、このときの$p_1\ p_2\cdots\ p_k$は、

$p_i=\frac{1}{k}$ $(i=1,2,\cdots,k)$ となる。

この$x_1,\ x_2,\cdots$ をϕに代入すれば、その値は $\frac{1}{4k}$ となる。□（定理1）

【系1】正確なコインをn回投げてn個の 0,1 の系列（i.e. $\{0,1\}^n$の要素$(a_1a_2\cdots a_n)$）を得て、擬似一様乱数をつくる場合、距離d_1を小さくするという意味で、すなわち分布関数をなるべく一様分布のものに近づけるという意味で最適な方法の一つは、$(0.\ a_1a_2\cdots a_n1)_2$ とすることである。

ここで「$(\ )_2$」は2進数で解釈することを意味する。

このとき、$d_1(D,U)=\frac{1}{2^{n+2}}$ となる。

□（系1）

例えばn＝3 回のコインを投げて、表—1、裏—0 を対応させ、$\{0,1\}^3$の要素から、d_1を小さくするという意味で最適な一様擬似乱数のつくり方の一つは

$000\rightarrow(0.0001)_2$、$100\rightarrow(0.1001)_2$

$001\rightarrow(0.0011)_2$、$101\rightarrow(0.1011)_2$

$010\rightarrow(0.0101)_2$、$110\rightarrow(0.1101)_2$

$011\rightarrow(0.0111)_2$、$111\rightarrow(0.1111)_2$

とすることである。

そして、このとき $d_1(D,U)=\frac{1}{32}$ となる(Fig. 8)。

4．距離 d_2の意味での最適な一様分布近似

まず、距離d_2の意味での最適な一様分布近似方法について議論するのに必要な補題を用意する。

【補題B】$D=\begin{pmatrix} x_1 & x_2 & \cdots & x_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}$

$x_{r-1}<a\leqq x_r<x_s\leqq b<x_{s+1}$ という区間[a,b]を考える（Fig.5）。この区間の集合をIとする。

（i.e. $I=\{[a,b];\ x_{r-1}<a\leqq x_r<x_s\leqq b<x_{s+1}\}$

また、$v\overset{def}{=}\sum_{i=r}^{s} p_i$ とする。

このとき、

$sup\{d_{[a,b]}(D,U);[a,b]\in I\}=$
$max\{|v-(x_s-x_r)|,\ |v-(x_{s+1}-x_{r-1})|\}$
となる。

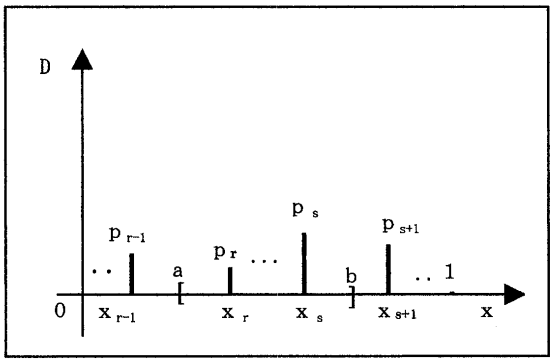

Fig.5　D and the Interval [a,b]

＜証明＞
$P_U[[a,b]]=b-a=t$ とすると、tは、
$x_s-x_r\leqq t<x_{s+1}-x_{r-1}$ の範囲を動く。
（1）$v\leqq(x_s-x_r)$ の場合
$d_{[a,b]}(D,U)=|P_D[[a,b]]-P_U[[a,b]]|$
$=t-v$　である。
この場合、
$sup\{d_{[a,b]}(D,U);[a,b]\in I\}=(x_{s+1}-x_{r-1})-v$
$=max\{|v-(x_s-x_r)|,\ |v-(x_{s+1}-x_{r-1})|\}$
（2）$v>(x_s-x_r)$ の場合　$d_{[a,b]}(D,U)=|v-t|$
（2－1）$(v>(x_s-x_r))\wedge(v-(x_{s+1}-x_{r-1})\geqq 0)$
の場合（このときは $d_{[a,b]}(D,U)=v-t$
となる。）
$sup\{d_{[a,b]}(D,U);[a,b]\in I\}=v-(x_s-x_r)$
$=max\{|v-(x_s-x_r)|,\ |v-(x_{s+1}-x_{r-1})|\}$
（2－2）$(v>(x_s-x_r))\wedge(v-(x_{s+1}-x_{r-1})<0)$
の場合
この場合の様子はFig.6にあるが、この図より、
どのような場合も、$sup\{d_{[a,b]}(D,U);[a,b]\in I\}$
$=max\{|v-(x_s-x_r)|,\ |v-(x_{s+1}-x_{r-1})|\}$
であることがわかる。□（補題B）

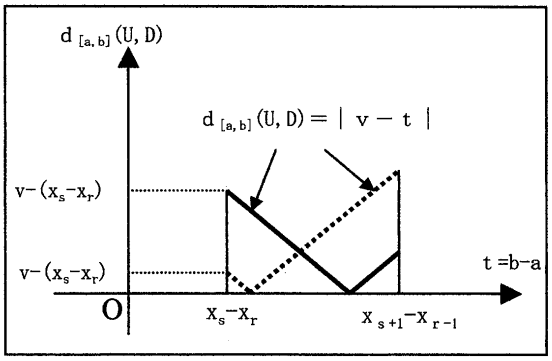

Fig.6　The Aspect of $d_{[a,b]}(U,D)$　(2 Cases)

補題Bで、特に　r＝sの場合は、
$sup\{d_{[a,b]}(D,U);[a,b]\in I\}$
$=max\{p_r,\ |p_r-(x_{r+1}-x_{r-1})|\}$ となる。

【定理2】$d_2(D,U)$を最小にする分布

$D_{2min}=\begin{pmatrix} x_1 & x_2 & \cdots & x_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}$ は、

$p_i=\frac{1}{k}$　$(i=1,2,\cdots,k)$、
x_i は $x_1\leqq\frac{1}{k}$、$1-x_k\leqq\frac{1}{k}$

$x_i-x_{i-1}\leqq\frac{1}{k}$　$(i=2,\cdots,k)$、
という分布である。
（この p_i と x_i $(i=1,2,\cdots,k)$ に関する条件を以下では「条件A」とよぶことにする。）
また、このときの $d_2(D_{2min},U)=\frac{1}{k}$　となる。
＜証明＞
まず、$d_2(D,U)$を最小にする分布 D_{2min} においては、その距離 $d_2(D_{2min},U)$ は、$d_2(D_{2min},U)\geqq\frac{1}{k}$　である。
なぜなら、
$d_2(D,U)=\underset{[a,b]\subset[0,1]}{sup}\{d_{[a,b]}(D,U)\}\geqq max\{p_1,p_2,..,p_k\}$
となる。これは各 x_i について
$a=x_i-\varepsilon<x_i<x_i+\varepsilon=b$ なる区間[a,b]で $\varepsilon\to 0$ とし $|P_D[a,b]]-P_U[[a,b]]|$ を考えれば明らかである。
そうすると、$d_2(D_{2min},U)<\frac{1}{k}$ ではあり得ないことがわかる（∵「$\forall i,\ p_i<\frac{1}{k}$」は $\Sigma p_i=1$ に矛盾する）。
したがって、　$d_2(D,U)=\frac{1}{k}$ となる分布Dがあれば、それが $d_2(D,U)$ を最小にする分布である。

「条件A⇒$d_2(D,U)$を最小にする」を証明する。
条件Aが成り立つとき、どのような区間[a,b]⊂[0,1]を考えても、$d_{[a,b]}(D,U)\leqq\frac{1}{k}$ であることがわかる。
なぜなら、$\exists x_i;\ x_i\in[a,b]$ の場合は

$a \leqq x_r < .. < x_s \leqq b$ として考えると（Fig.7）、補題Bより、$d_{[a,b]}(D,U) \leqq$

$\max\{|v-(x_s-x_r)|, |v-(x_{s+1}-x_{r-1})|\}$

$\leqq \max\{\frac{m}{k}-\frac{m-1}{k}, \frac{m+1}{k}-\frac{m}{k}\} = \frac{1}{k}$ となる（ここで、$m=|\{x_i ; x_i \in [a,b]\}|=s-r+1$）。

また、$\sim(\exists x_i ; x_i \in [a,b])$ の場合は、

$d_{[a,b]}(D,U)=b-a$

$\leqq \max\{|x_i - x_{i-1}| ; i=1,2,..,k+1\}$

（ここでは、$x_0=0$，$x_{k+1}=1$ としておく）

である。したがって仮定より $d_{[a,b]}(D,U) \leqq \frac{1}{k}$ となる。

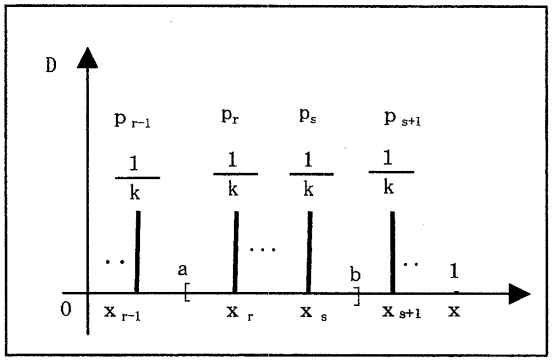

Fig.7 The Aspect of D_{2min}

これまでの議論により、定理2の条件Aを満たす分布は $d_2(D,U)=\frac{1}{k}$ となり、$d_2(D,U)$ を最小にする分布であることが明らかになった。

「$d_2(D,U)$ を最小にする分布 D⇒定理2の条件Aを満たしている」ことを証明する。

$\exists i ; x_i - x_{i-1} > \frac{1}{k} \quad \vee \exists i ; p_i > \frac{1}{k}$

であれば、$d_2(D,U) > \frac{1}{k}$ となることは明らかである。 □（定理2）

定理2より、$p_i=\frac{1}{k}$, $x_i=\frac{i-1}{k}$ $(i=1,2,\cdots,k)$ という分布 D_1 は $d_2(D,U)$ を最小にする分布の一つである。

また、$p_i=\frac{1}{k}$、$x_i=\frac{i}{k}$ $(i=1,2,\cdots,k)$ という分布 D_2 も $d_2(D,U)$ を最小にする分布の一つである。

このように $d_2(D,U)$ を最小にする分布はユニークには決まらない。

【定理3】 $d_1(D,U)$ を最小にする離散分布Dは $d_2(D,U)$ を最小にする。

＜証明＞定理1、定理2より明らか。□（定理3）

Fig.8に、3回のコイン投げ（k=8の離散分布D）から一様乱数を得る場合の d_1 の意味で（したがって d_2 の意味でも）最適な方法の、密度関数、分布関数を示す。

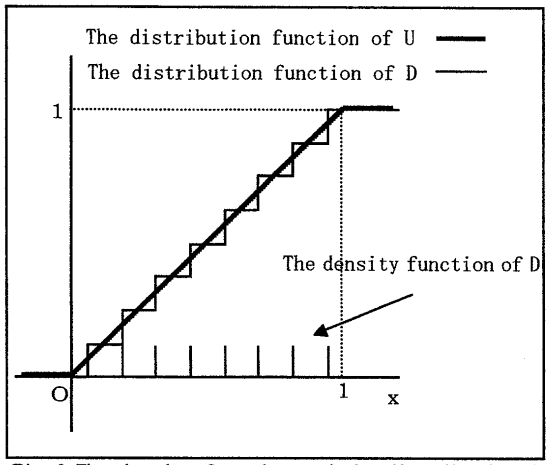

Fig.8 The density function and the distribution function of the Optimal Method for Generating Random numbers from 3-honest coins tossing

5．結言および今後の課題

理想的な離散分布が与えられたとき、それから[0,1]上の一様分布に最も近い分布を生成する方法について論じた。今後の課題としては、情報の変換、例えば、前述の、一様分布にしたがう乱数を用いるモンテカルロ法等を行う際に、乱数情報の数値近似情報への変換等についてその性質を調べることがあげられる。その際に、本論文で調べた離散分布による一様分布の最適な近似に関する知識が必要になると考えられる。

参考文献

1) アブラムソン（宮川訳）：情報理論入門，好学社，1969年
2) 上田顕、コンピュータシミュレーション，朝倉書店，1990年
3) 田村進一：情報工学基礎論，培風館，1992年
4) 伏見正則：確率的方法とシミュレーション，岩波書店（岩波講座 応用数学）pp1-22，1994年
5) 三根 久：モンテカルロ法・シミュレーション、pp25-48、コロナ社、1994年
6) 宮武 修、脇本和昌：乱数とモンテカルロ法、森北出版，1978年
7) 脇本和昌：乱数の知識、森北出版 1970年
8) 涌井良幸、涌井貞美：パソコンで遊ぶ数学実験、pp28-41、講談社ブルーバックス 2003年