札幌学院大学総合研究所紀要（2024）第11巻 1-7　　［研究ノート］

疑似乱数の一様性の尤度による評価

中村　永友[1]　　土屋　高宏[2]

要　旨

本論文は区間（0,1）内の点列の一様性を確認する１つの方法として，ベータ分布の平均対数尤度を通した評価を試みる．さらに一様乱数の性質を基礎とする理想的な点列を提示し，その特性に言及する．さらに，平均対数尤度や尤度関数による一様分布の評価は，確率分布の特性を反映しないので，ベータ分布のカルバック＝ライブラー情報量により，データの一様性の検討を行う．このことにより少数のデータの一様性の評価が可能になる．

キーワード：一様分布，平均対数尤度，対数尤度，準一様乱数，順序統計量，カルバック＝ライブラー情報量，理想点列，ベータ分布

1　はじめに

一様分布にしたがうと考えられるデータがあり，その一様性を尤度を使って検討をしたい．しかし，一様分布の特殊性から（対数）尤度関数ではその評価ができない．これを克服する方法の１つとして，一様分布に従うデータの順序統計量の差分がベータ分布にしたがうという性質を使って，ベータ分布の対数尤度を通した一様性の検討が可能になる．

疑似一様乱数などの点列の乱数性を確認するには，無作為性と一様性をそれぞれ検証する必要があるが，本稿では一様性のみに注目する．一様性を確認するには，多数（数100程度以上）のデータに対して，適合度検定や KS 法等で均等に分布しているかを統計的な手続きで検証ができる．例えば NIST SP800-22（NIST Test Suite, 2010; Andrew *et al*., 2010）や TestU01（Pierre & Richard, 2007）などの一連の検定の手続きもある程度有効である．しかし，尤度による評価が可能になれば, 少数のデータに対する検討が促進される．

一方，データが与えらたとき，そのデータが指定された確率分布に従うと考えられるか否かを判断する方法として，尤度を使うことができる（坂元他，1983）．つまり指定した確率分布の平均対数尤度の値と，データから得られる対数尤度をカルバック＝ライブラー情報量で比較ができる．この方法の利点は，ごく少数のデータでもその確率分布への適合性を判断できることである．

本報告は一様乱数列の持つ統計的な性質を利用した一様乱数の生成法を提案する一連の研究（中村・土屋，2021，2023）の続報である．既存のものと生成過程が異なる疑似乱数を提案し，これにより得られた一様乱数を準一様乱数（quasi-unifom random number）と呼んだ．本報告は，準一様乱数の持つ統計的な性質の一面を明らかにすることと，一様分布にしたがうデータを尤度による評価を試みる．特に断りがない場合は，疑似乱数を単に「乱数」と表記する．

以下，第２節ではカルバック＝ライブラー情報量の確認をする．第３節では一様分布における尤度の特性であり，統計的推定における不具合を指摘する．第４節ではベータ分布の平均対数尤度の算出，第５節で理想点列としてのギャップ点列，ベータ点列，準ギャップ点列について説明し，第６節では理想的点列の対数尤度の評価をする．

2　カルバック＝ライブラー情報量

カルバック＝ライブラー情報量（以下，KL 情報量；Kullback and Leibler, 1951）とは２つの確率分布の差異を計る尺度である．確率分布が連続型のとき，２つ

[1] 札幌学院大学　経済経営学部；nagatomo@sgu.ac.jp.
[2] 城西大学　理学部；takahiro@josai.ac.jp.

の確率分布が P と Q，その確率密度関数が $p(x)$ と $q(x)$ のとき，以下のように定義される：

$$D_{\mathrm{KL}}(P\|Q)=\int_{-\infty}^{\infty}p(x)\log\frac{p(x)}{q(x)}dx=E_P\left[\log\frac{p(x)}{q(x)}\right]$$
$$=E_P[\log p(x)]-E_P[\log q(x)].$$

E_Pは確率分布 P を基準として期待値をとるということである．すなわち P の視点から Q を測っていることになる．一方，$E_P[\log p(x)]$は基準となる確率分布の平均対数尤度である．情報量規準では，データを生成する真の確率分布を P として，これを未知として扱う．想定する確率分布モデルを Q として，得られたデータ$\{x_1,\dots,x_n\}$を元に対数尤度（の $1/n$）

$$E_P[\log q(x)]=\frac{1}{n}\sum_{i=1}^{n}\log q(x_i|\theta)$$

を最大にする θ を求めることで，より P に近づけることをしている．KL 情報量は P と Q が同じ確率分布の時は 0 になるという性質を持つので，対数尤度は平均対数尤度の推定値ということになる．

ここで，真の確率分布 P を既知として，仮定する確率分布モデル Q も P と同じ確率分布とする．この条件下で対数尤度が平均対数尤度の推定値であることについて考える．

例 1　P と Q を正規分布 $N(0,1)$として，1 組のデータセット $X=\{x_1,\dots,x_n\}$を $N(0,1)$から生成し，対数尤度 $L(X)=\frac{1}{n}\sum_{i=1}^{n}\log f(x_i)$ を求める．この $L(X)$ は $N(0,1)$の平均対数尤度の推定値ではあるが，多くのデータセットを生成して，$E[L(X)]$を求めることで，平均対数尤度と一致することが確認できる．1 組の X からの $L(X)$ も推定値ではあるが，真の確率分布 P が既知なので，その平均対数尤度との差分で推定値の近さを測ることが出来る．このとき，尤度の視点から，その差分が 0 に近いときは，データセット X が真の確率分布 P からの非常に良い無作為標本であると言える．ここで注意することは，$N(0,1)$からの無作為標本を前提としていることである．無作為標本でも低い確率で偏りが生じ，そのような場合の対数尤度は平均対数尤度からも離れた値になる．　□

標準正規分布から発生した乱数が 2 組あり，それぞれ対数尤度を計算したところ，-1.450と-1.351であった．$N(0,1)$の平均対数尤度は-1.41894なので，この値により近い前者がより正規分布に近いデータセット（より正規分布している）と考えることができる．つまり正規分布の状態をより良く表した無作為標本であると言える．　□

例 2　真の確率分布をベータ分布 Beta（$1,n$）とし，これより無作為標本を抽出して，対数尤度で平均対数尤度を推定した数値実験を表 1 に示す．n は抽出した標本の数，繰り返し抽出を $S=100{,}000$回として対数尤度の期待値を求めた．また同時にその標準偏差や経験分布からパーセンタイルを求めた．対数尤度関数が平均対数尤度の良い推定値であることがわかる．　□

3　一様分布と尤度

確率分布とそのパラメータ値が特定されれば，平均対数尤度は理論的に求めることができ，得られたデータの対数尤度から真の確率分布 P に対する適合性を検討することが可能である．しかし，一様分布は特殊であり，例外的である．

系 1（一様分布の平均対数尤度）　区間$(0,1)$で定義される確率密度関数が $f(x)=1$ の一様分布の平均対数尤度は 0 である．

証明

$$\int_0^1 f(x)\log f(x)dx=\int_0^1 1\cdot\log 1dx=0.\qquad\square$$

系 2（一様分布の対数尤度）　区間$(0,1)$内でデータがどんな分布をしていても，一様分布 $U(0,1)$を仮定した対数尤度の値は 0 である．

証明　区間$(0,1)$から生成した一様乱数のデータセット $X=\{x_1,x_2,\dots,x_n\}$がある．$U(0,1)$の密度関数 $f(x)=1$ として対数尤度を求める：

$$L=\sum\log f(x_i)=\sum\log 1=0.\qquad\square$$

系 2 より，真の密度関数が一様分布を前提とした場合にはどんなデータに対しても対数尤度が 0 になる．さらに平均対数尤度の値もである．つまり，一様分布の対数尤度はデータの状況を反映できないということである．これは他の確率分布にない特殊性である．

適合度検定のような方法ではなく，データの一様性を尤度で判定するにはどのような方法があるか．1 つの方法として，一様分布しているデータ$\{u_1,u_2,\dots,u_n\}$において，順序統計量 $u_{(k)}$ ($k=1,\dots,n;\ u_0=0,u_{n+1}=1$) の差分 $b_{(k+1)}=u_{(k+1)}-u_{(k)}$がベータ分布 Beta$(1,n)$に従うことから（David & Nagaraja, 2003），この関係を

表 1：ベータ分布 Beta（1, n）の平均対数尤度と対数尤度の分位数（S＝100,000）

n	平均対数尤度	KL情報量	対数尤度のS.D.	パーセンタイル 0.01	0.025	0.05	0.25	0.50	0.75	0.95	0.975	0.99
5	0.809	−0.00106	0.327	−0.139	0.054	0.210	0.621	0.854	1.047	1.263	1.320	1.375
10	1.406	−0.00073	0.271	0.657	0.799	0.916	1.239	1.430	1.598	1.800	1.855	1.914
20	2.046	0.00067	0.207	1.498	1.600	1.681	1.914	2.061	2.191	2.359	2.407	2.460
50	2.932	0.00054	0.137	2.585	2.644	2.695	2.843	2.938	3.027	3.147	3.183	3.224
100	3.615	0.00026	0.098	3.373	3.414	3.448	3.551	3.618	3.683	3.770	3.798	3.830
200	4.303	−0.00023	0.070	4.134	4.163	4.186	4.257	4.305	4.352	4.416	4.436	4.459
500	5.217	0.00012	0.045	5.109	5.127	5.142	5.187	5.217	5.247	5.289	5.303	5.318
1,000	5.909	0.00009	0.032	5.834	5.846	5.856	5.887	5.909	5.930	5.960	5.970	5.981
10,000	8.210	0.00006	0.010	8.187	8.191	8.194	8.204	8.210	8.217	8.227	8.230	8.234

使って，ベータ分布を通した尤度が考えられる．つまり，順序統計量の差分がベータ分布していれば，元のデータは一様分布していることになる．

4　ベータ分布の平均対数尤度と対数尤度

ベータ分布 Beta$(1,n)$の平均対数尤度は次式である：

$$f(x|n)=n(1-x)^{n-1},$$
$$E[\log f(x|n)]=\int_0^1 f(x|n)\log f(x|n)dx$$
$$=\frac{1}{n}+\log n-1.$$

データ $\boldsymbol{B}=\{b_1, b_2, ..., b_{n+1}\}$が与えられたときのベータ分布の対数尤度：

$$L(\boldsymbol{B})=\sum_{i=1}^{n+1}\log\{n(1-b_i)^{n-1}\}$$
$$=(n+1)\log n+(n-1)\sum_{i=1}^{n+1}\log(1-b_i).$$

手元のデータに対してベータ分布への当てはまりの程度を対数尤度の値で確認すれば，間接的にデータの一様性をカルバック＝ライブラー情報量で確認できる．あるいは対数尤度の（シミュレーションなどにより）分布がわかれば仮説検定のようなことができる．

系 3　区間$(0,1)$内の n 個からなる任意のデータ点列 X を順序統計量に変換し，隣り合う 2 つの値の差分に対してベータ分布 Beta$(1,n)$の対数尤度で X の一様性の検討が可能である．　□

これは定理 A2(付録 A)と系 2 により明らかである．

5　理想点列

中村・土屋（2021，2023）により提示された準一様乱数の種となる 3 つの点列，ギャップ点列，ベータ点列，準ベータ点列を**理想点列**（ideal sequence）と呼ぶ．これらはいずれも区間$(0,1)$内の一様乱数の理論的背景により導出されている．その概要を以下に示す．理想点列から乱数を生成する方法は付録 B に示す．

5.1　ギャップ点列

ギャップ点列 $E[g_{(k)}]$は区間$[0,1]$上の一様分布から抽出した n 個の標本を順序統計量にして，隣り合う値の差分（ギャップ）の期待値で与えられる．それは以下の(1)式で得られる：

$$G(k)=E[g_{(k)}],\ (k=1,2,...,n).$$

（i）区間$(0,1)$から独立に抽出した一様乱数を$\{x_1, x_2, ..., x_n\}$とし，順序統計量$\{x_{(1)}, x_{(2)}, ..., x_{(n)}\}$とする．また，$x_{(0)}=0, x_{(n+1)}=1$とする．

（ii）隣り合う 2 つの差分（ギャップ）$g_{i+1}=x_{(i+1)}-x_{(i)}$，（$i=0, ..., n$）を求める．$g_i$ はベータ分布 Beta（$1, n$）にしたがう（David & Nagaraja，2003）．

（iii）g_iを順序統計量 $g_{(i)}$にする．

（iv）k 番目の順序統計量 $g_{(k)}$の期待値は，

$$E[g_{(k)}]=\frac{1}{n}\left\{\sum_{i=1}^{n}\frac{1}{i}-\sum_{j=1}^{n-k}\frac{1}{j}\right\}=\frac{1}{n}\sum_{j=n-k+1}^{n}\frac{1}{j} \qquad (1)$$

で与えられる（中村・土屋，2023）．

5.2　ベータ点列の導出

区間$[0,1]$上の一様分布から抽出した n 個の標本の順序統計量の確率分布を元に作られるのが，**ベータ点列** $E[b_{(k)}]$ $(k=1, 2, ..., m)$である．それは次の方法に

よって得られる：

（i）ベータ分布 Beta$(1, n)$ から独立に $m(=n+1)$ 個乱数を抽出し，これを b_i とする．

（ii）これを順序統計量 $b_{(i)}$ にする $(i=1, ..., m)$．

（iii）$E[b_{(i)}]$ を求めるために，$b_{(i)}$ の確率分布を求める．

順序統計量にたいする一般的な分布関数は

$$F_k(x)=\sum_{i=k}^{n}\binom{n}{i}\cdot\{F_X(x)\}^i\cdot\{1-F_X(x)\}^{n-i}$$

であるので（David & Nagaraja，2003），ここではベータ分布 Beta$(1, n)$ の分布関数と密度関数

$$F_X(x)=1-(1-x)^n,\ f_X(x)=n(1-x)^{n-1}$$

を用いる．また，$E[b_{(i)}]=E[f_k(x)]$ である．

（iv）k 番目の順序統計量の密度関数は次式である：

$$f_k(x)=\frac{m!}{(k-1)!(m-k)!}\left\{1-(1-x)^n\right\}^{k-1}\times n(1-x)^{n(m-k+1)-1}.$$

（v）k 番目の順序統計量の密度関数の期待値は次式である：

$$E[b_{(k)}]=E[f_k(x)]=\frac{n\cdot m!}{(m-k)!}\sum_{j=0}^{k-1}\frac{(-1)^{k-1-j}}{(k-1-j)!j!}\times\frac{1}{(nm-nj+1)(nm-nj)}.$$

5.3　準ベータ点列：シミュレーションによる導出

準ベータ点列（規格化したベータ点列）は，順序統計量の差分がベータ分布したがうことを利用して，総和が1になるように調整しながらその期待値を求めるものである．以下の手順で得られる：

（i）ベータ分布 Beta（$1, n$）から独立に $n=m+1$ 個乱数を抽出し，これを c_i' とする．

（ii）$C=\sum c_i'$ として，$c_i=c_i'/C$ を求める．

（iii）これを順序統計量にして $c_{(i)}$ とする（$i=1, ..., m$）．

（iv）$E[c_{(i)}]$ をシミュレーションで求める．

6　ベータ分布の平均対数尤度との比較

理想点列により得られる対数尤度が平均対数尤度の推定値としてどのような特性があるか確認する. 3種の理想点列を$\{g_1, g_2, ..., g_{n+1}\}$, $\{b_1, b_2, ..., b_{n+1}\}$, $\{c_1, c_2, ..., c_{n+1}\}$として, これらのベータ分布 Beta$(1, n)$ に対する対数尤度の値を求める. これらが順序統計量であるか否かは対数尤度を計算する上では関係ない. その結果を表2に示す.

3種類の点列で，カルバック＝ライブラー情報量からはいずれも若干の過大推定ではあるが，平均対数尤度の値と一致しているとみなしても良いものが得られた．推定値としてはかなり良好である．これらの中でもっとも近いのがギャップ点列である．ギャップ点列はベータ点列とは同様な理論によるものだが，ベータ分布ではない異なる確率分布を考えるのが相当と考えられる．しかし，対数尤度の値はベータ点列よりも優れている．ベータ点列は順序統計量の期待値を求めているが，これは非対称な分布であることから，中央値などを考えた方がよいのかもしれない．

以上の結果から，理想点列は平均対数尤度の非常に良い推定値であり，確率分布を代表する尤もらしいデータ（点列）である．理想点列を持つ順序統計量は，好ましい一様乱数列であることがわかる．

既知の確率分布に対する平均対数尤度（理論式）は分布固有のものであるが，理論値（スカラー値）は固有ではない．３種の理想点列は総和が１で，いずれも一様乱数の関係理論から導出されたという背景を持つ．総和が１であるという条件のみから，平均対数尤度と同じ値を持つ点列を作ることはできるが（付録B），その点列が好ましい一様乱数列となるには，何らかの一様分布由来であることの証明が必要である．これに関することは今後の研究課題である．

表１に経験分布から期待値やパーセンタイルを示したが，１万回単位の反復実験では，真の値と期待値が一致するのは，ある意味あたりまえである．しかし表２は理想点列を１組のみで，平均対数尤度をピタリと推定しているのである．

対数尤度が漸近正規性を持つことから，表３に対数尤度の期待値と分散を元として，理想点列で得られた対数尤度の値の絶対値以内の正規分布の確率を求めた結果を示す．これらの確率が小さいほど平均対数尤度に近い値が推定されていることを表す．

7　おわりに

本報告は，一様分布の特殊性を示すと共に，疑似一様乱数の一様性をベータ分布を通して示す方法を提示した．中村・土屋（2021，2023）が提案した準一様乱数（理想点列）は，平均対数尤度の推定値という点ではベータ分布を代表する点列であることが判った．一般の確率分布でそのような点列が得られるか，また，理想点列と主要点（principal points; Flury，1990）と

表 2：ベータ分布 B（1，n）の平均対数尤度，理想点列の対数尤度（LL）と平均対数尤度の差（KL 情報量）

n	平均対数尤度	ギャップ点列		ベータ点列		準ベータ点列	
	ELL	*LL*	*KL* 情報量	*LL*	*KL* 情報量	*LL*	*KL* 情報量
5	0.809438	0.827107	−0.0176693	0.839731	−0.0302929	0.837271	−0.0278332
10	1.40259	1.40963	−0.00704452	1.41454	−0.0119520	1.41382	−0.0112309
20	2.04573	2.04819	−0.00246156	2.04980	−0.00406411	2.04976	−0.00403111
50	2.93202	2.93257	−0.000549286	2.93289	−0.000864202	2.93289	−0.00087140
75	3.33082	3.33110	−0.000275798	3.33125	−0.000424103	3.33124	−0.000421817
100	3.61517	3.61534	−0.000168000	3.61542	−0.000254187	3.61543	−0.000262320
200	4.30332	4.30337	−0.0000499251	4.30339	−0.0000727384	4.30339	−0.0000767748
500	5.21661	5.21662	-9.71618×10^{-6}	5.21662	−0.0000135310	5.21662	−0.0000132655
1,000	5.90876	5.90876	-2.76322×10^{-6}	5.90876	-3.73521×10^{-6}	5.90876	-4.17563×10^{-6}
10,000	8.21044	8.21044	-3.89759×10^{-8}	8.21044	-4.89282×10^{-8}	8.21044	-5.57222×10^{-8}

表 3：理想点列の対数尤度の絶対値以内の確率

n	ギャップ点列	ベータ点列	準ベータ点列
5	0.1340	0.2276	0.2096
10	0.0775	0.1310	0.1232
20	0.0464	0.0766	0.0760
50	0.0237	0.0373	0.0376
75	0.0175	0.0268	0.0276
100	0.0141	0.0213	0.0220
200	0.0181	0.0118	0.0125
500	0.0038	0.0053	0.0052
1,000	0.0022	0.0029	0.0033
10,000	0.0001	0.0004	0.0004

の関係など，興味のあるところである．理想点列の対数尤度が良い推定値であることは明確だが，若干の過大推定気味なので，なぜぴったり一致しないのかなど，今後の研究課題としたい．

謝辞　本研究は2023年度札幌学院大学研究奨励金SGU-BG2023-03の助成を受けた．

参考文献

[1] Andrew R., Juan S., James N., Miles S., Elaine B., Stefan L., Mark L., Mark V., David B., Alan H., James D. and San V.（2010）. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, Special Publication 800-22 Revision 1a, NIST（National Insititute of Standatds and Technology）, https://nvlpubs. nist. gov/nistpubs/legacy/sp/nistspecial publication800-22r1a.pdf.

[2] NIST Test Suite（2010）. OpenSSL. Openssl implementation from http://www.openssl.com/, http://csrc.nist.gov/groups/ST/toolkit/rng/.

[3] David, H. A. and Nagaraja, H. N.（2003）. Order Statistics, Third Edition, Wiley.

[4] Flury, B.A.（1990）. Principal points. Biometrika, 77, 1, 33-41.

[5] Kullback, S., and Leibler, R. A.,（1951）. On information and sufficiency, Annals of Mathematical Statistics, 22, 79-86.

[6] 中村永友・土屋高宏（2021）. 擬似的な一様乱数とベータ分布, 札幌学院大学 総合研究所紀要（情報科学）, 8, 57-65.

[7] 中村永友・土屋高宏（2023）. 一様乱数列生成のためのベータ分布を基礎とする点列, 札幌学院大学 総合研究所紀要（情報科学）, 10, 1-8.

[8] Pierre L. and Richard S.（2007）. TestU01: A C library for empirical testing of random number generators, ACM Transactions on Mathematical Software, 33, August 2007 Article No.: 22, https://doi.org/10.1145/1268776.1268777, http://simul.iro.umontreal.ca/testu01/ tu01.html.

[9] 坂元慶行・石黒真木夫・北川源四郎（1983）. 情報量統計学, 共立出版.

付録

A　関係する定理や定義

最初にいくつかの定義と定理を示す．

定義 A1（順序統計量）　区間$(0,1)$の一様乱数$\{x_1, x_2, \ldots, x_n\}$の順序統計量を$\{x_{(1)}, x_{(2)}, \ldots, x_{(n)}\}$で表す．

定義 A2（ベータ分布）　ベータ分布 $\mathrm{Beta}(\alpha,\beta)$の定義：

$$f(x\,|\,\alpha,\beta)=\frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha,\beta)},\ (0<x<1).$$

$\mathrm{B}(\cdot,\cdot)$はベータ関数，期待値と分散は次の通り：

$$E[X]=\frac{\alpha}{\alpha+\beta},\ \ V[X]=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

定理 A1（順序統計量の分布）　区間$(0,1)$における一様乱数に対する順序統計量$u=x_{(r)}$は，次の確率分布にしたがう（David & Nagaraja，2003）：

$$f(u)=\frac{1}{\mathrm{B}(r,n-r)}u^{r-1}(1-u)^{n-r-1},\ 0\leq u\leq 1.$$

期待値と分散は次の通り：

$$E(u)=\frac{r}{n},\ E(u^2)=\frac{r(r+1)}{n(n+1)},\ V(u)=\frac{r(n-r)}{n^2(n+1)}.$$

定理 A2（差の分布）　一様分布にしたがう標本の順序統計量の差分$x_{(k)}-x_{(j)}(k>j)$は，ベータ分布 Beta$(k-j, n-k+j+1)$にしたがう（$1\leq j<k\leq n$; David & Nagaraja，2003）.

系 A1（隣接順序統計量の差の分布）　定理 A2より順序統計量の差分$x_{(i+1)}-x_{(i)}$はベータ分布 Beta$(1,n)$にしたがう.

系 A2　ベータ分布 Beta$(1,n)$の密度関数：

$$f(x\,|\,1,n)=\frac{x^0(1-x)^{n-1}}{\mathrm{B}(1,n)}=n(1-x)^{n-1},\ (0<x<1).$$

系 A3（ベータ分布にしたがう確率変数の和）　$n+1$個のベータ分布 Beta$(1,n)$にしたがう乱数の和bは，$b\fallingdotseq 1$である.

証明　Beta$(1,n)$の期待値が$\frac{1}{n+1}$であるので，$n+1$個のベータ分布にしたがう乱数の和は１である：

$$E\left[\sum_{i=1}^{n+1}X_i\right]=\sum_{i=1}^{n+1}E[X_i]=\sum_{i=1}^{n+1}\frac{1}{n+1}=1.$$

系 A4　Xが一様乱数列であることの必要十分条件は，Xの順序統計量の差がベータ分布であることである.

系 A5　一様乱数の順序統計量の差がベータ分布にしたがうことを利用すると，ベータ分布に対する適合度検定で間接的に一様性の検定が可能である.

B　理想点列から一様乱数を作る

理想点列を$\{r_1, r_2, ..., r_{n+1}\}$とする．これをシャッフルして（あるいは非復元抽出をして）改めて$\{r_1, r_2, ..., r_{n+1}\}$と並べる．これを累積する：

$$x_1=r_1, x_2=x_1+r_2, ..., x_i=x_{i-1}+r_i, ..., x_n=x_{n-1}+r_n.$$

これにより得られた$\{x_1, x_2, ..., x_n\}$が区間$(0,1)$内の一様乱数となる.

C　平均対数尤度の値と同じ値になる点列

ベータ分布 Beta$(1, n)$の平均対数尤度$\frac{1}{n}+\log n-1$で，$n=10$のときの値は1.402585である．この値と対数尤度が等しくなる点列を

{0.0656215, 0.0656215, 0.0656215, 0.0656215, 0.0656215, 0.0656215, 0.0656215, 0.0656215, 0.0656215, 0.0656215, 0.343785}

と作ることが出来る．この点列の総和は１，10個の0.0656215と１個の0.343785からなる．あくまでも１例であり他にも作成可能である．これらを元にした点列を図に示すが，均等間隔に並んだ10個と離れた１つの点列を一様乱数列とみなせなくはないだろうが，かなり困難ではないだろうか.

図 C1：平均対数尤度と同じ値となる点列

The Proceedings of the Research Institute of Sapporo Gakuin University Vol.11, 1–7 (2024)

An Evaluation Method for Uniformity of Pseudo-Random Numbers via Log-Likelihood

Nagatomo NAKAMURA[1]
and
Takahiro TSUCHIYA[2]

Abstract

This article attempts to evaluate a method for checking the uniformity of point sequences sampled within the interval (0, 1) using the expected log-likelihood of the beta distribution. Furthermore, we present the ideal point sequence based on the properties of the uniform random numbers and discuss its properties. Since evaluation using the expected log-likelihood or likelihood function of the uniform distribution does not reflect the characteristics of the probability distribution, the uniformity is examined using the Kullback=Leibler information content of the beta distribution. This makes it possible to evaluate the uniformity of a small number of data.

Keywords: Beta Distribution, Expected Log-Likelihood, Log-Likeihood, Ideal Sequence, Kullback=Leibler Information, Order Statistics, Quasi-Random Number, Uniform Distribution.

[1] Department of Economics, Sapporo Gakuiun University; nagatomo@sgu.ac.jp.
[2] Department of Mathematics, Josai University; takahiro@josai.ac.jp.