数理解析研究所講究録
第498巻 1983年 13-22

乱数の発生・変換について

統計数理研　仁木　直人（Naoto Niki）

§0. はじめに

統計的シミュレーションやモンテカルロ計算などに用いられる乱数およびその発生法については、各種の（ときには相反する）性質が要求される。ここでは、半ば定形化した議論とは異なり、乱数の発生・変換法に関してのかなり厳しい見方を提示しようと思う。一応、議論の対象は算術乱数とするが、物理乱数についても、対比の意味で、簡単に触れる。

§1. 乱数の使用個数

まず考慮しなければならないのは、全部で（同様な実験を繰返すならば、その合計で）どれくらいの個数の乱数を用いるのか、ということである。

使用する乱数の個数があまり多くない場合（問題にもよるが10万個のオーダ以下と考えれば良いだろう）には、偶然による変動が大きいので、敢て「問題とすべきことは殆どない」と謂おう。つまり、余程ひどい発生・変換法でなければ、少々発生速度が遅くても、乱数の性質に多少の不安があっても、何で

もよいわけである。もし手元に発生・変換のためのプログラムがあれば、それを使えば良いし、もしなければ、「簡単に作れるもの」を捜せば事足りる。汎用の乱数パッケージが使えれば理想的である。

問題は、使用する乱数の個数が非常に多い場合（１０００万個程度以上）である。勿論、発生・変換の高速性や経済性（変換の元となる乱数を平均何個必要とするか）も重要なファクタとなるが、謂わば二次的な事項であろう。忘れてならないことは、「多くの問題では、用いる乱数が『全く理想的な性質を有している』との前提の下で結論を導く」、ということである。特に、物理現象のシミュレーションでは、乱数の性質の良し悪しが明確に出易い。それゆえ、そこで用いる乱数は、発生・変換法に関する原理の妥当性や数学的厳密性を問われるばかりでなく、通常の検定より一段階厳しいチェックを必要とする。例えば、「検定のパワーの比較をモンテカルロ的に行なう」場合などを考えれば、その理由も理解されよう。以下の章では、使用個数の多い場合について、さらに詳しい議論を行なおう。

§２．算術乱数発生法（一様乱数）に関して

最もよく議論される性質は、高速性と可搬性（簡潔性、省空間性などを含む）であろう。この両者はしばしば背反する傾向があるが、どちらがより重要かと謂えば、迷わず「高速性」の方を採る。しかし、これらは所詮二次的な性質にすぎない。始

めに問われるべきは、「発生原理そのものの妥当性」である。

では、算術乱数の発生に当って、「原理の妥当性」は本当にあるのか？　『全く無い』のである。少なくとも、簡単な漸化式に従って決定論的に順次定まる値の列を「乱数列」と呼ぶのは、相当図々しい行為と謂えよう。ボロの出ない方がむしろ不思議である。我々は、このような認識を踏まえた上で、算術乱数の性質等の議論を行なう必要がある。

生成される算術一様乱数についてのまず第一の関心事は、周期の長さ、擬周期の存在とその影響、乱数として使用できる有効数字の桁数、それに周期について成り立つ各性質（適合性、独立性）などである。これらに関しての種々の議論は他に譲る（Knuth（1981）など）こととして、ここではあまり他で採上げられない点を指摘するに留める。それは簡単な変換に対する耐性である。例えば、（0，1）乱数は対数をとってから使われることが多いが、この場合は0に近い所の諸性質（有効数字など）が大きな問題となろう。

より実際に即した観点からは、「周期内の部分列の統計的性質」の方がはるかに重要である。そのチェックには、次の二方面からの実測が必須と思われる。

第一は、「実際に使われる程度の長さを持つ部分列に関し、その全体について求めた統計量及びそれに基づく推論」で、通

常の「乱数の検定」やコレログラムあるいはパワー・スペクトルを求めることなどがこれに当る。実際に使用する予定以外の数個の部分列に対しても試みるべきである。

第二は、「実際に使われる程度の長さを持つ部分列に関し、その短い構成単位ごとに求めた統計量の『分布』及びそれに基づく推論」である。表1はその一例で、代表的ないくつかの発生法について、16進乱数4096個ごとにその一様性の尺度であるχ^2値を算出し、そのχ^2値300000個の分布が自由度15のχ^2分布に従っているか検定を行なった結果である。m系列は短い構成単位ごとの性質があまり芳しくないようである。詳細は、仁木（1983）に発表予定である。

さらに「適用する問題に固有な使用法に対する耐性」についても慎重な検討を要する。既知の解を持つ類似した問題に適用してのチェックが可能ならば非常に好ましい。また、「常に2個を対にして使う」などの特徴がある場合には、「対（2次元乱数）としての性質」を集中的に調べるなどの『狙い打ち』も必要であろう。

物理乱数を使用すれば、発生原理のまともな分だけ、気苦労が少なくて済む。物理乱数の場合は、比較的単純な検定で不都合な性質が、もしあれば、明らかになると期待できるからである。物理乱数については、仁木（1980）に総合的報告がある。

§３．乱数の変換に関して

一様乱数の発生に関する事項に加えて、「経済性」「数学的厳密性」「特殊関数の使用の有無（可搬性の一種）」などが問題となるが、ここでは他で論じられることの少ない重要な二点について述べる。

第一は、「元になる乱数が離散値を取ることによる影響」である。変換後の乱数の厳密な分布、特に分布の裾に於ける影響を調べる必要がある（図１）。離散値の間隔を変えているに過ぎない「逆関数法」（図２）や単純な「envelope rejection」などでは、その影響が出やすい。

第二は、「変換後の乱数の性質が元になる乱数の性質から、直接あるいは間接的に、導出または推測できるかどうか」である。「直接導出」可能ならば理想的であるが、そこまで行かなくとも、「原理が単純で、homogeneous な変換法」が好ましいと謂えよう。それは、予測できないような変な性質を背負い込む恐れが少ないことと、元になる乱数の性質の内のどの部分を『狙い打ち』に調べればよいかが分かる点からである。

第一・第二の両方の観点のみからは、「単純棄却法」を採用することのメリットは大きい（Bowmanは「他の方法は信用しない」と言明している）。離散値の間隔を変えない上、元になる乱数を３個ずつ組合わせた３次元乱数の一様性および独立性が証明されれば、変換後の乱数の適合性および独立性が直ちに保証される。Niki（1979）による正規乱数発生法（図３）は、単

純棄却法の特長を損うことなく高速化を図った例である。

参考文献

Bowman, K. O. (1979) personal communication.

伏見正則、手塚集（1981）多次元分布が一様な擬似乱数列の生成法．応用統計学１０巻３号．

石田、佐藤、鈴木、下田、川瀬（1972）ダイオード・ノイズを利用した乱数発生装置．日立評論５４巻１０号．

Knuth, D. E. (1981) The art of computer programming. Vol. 2 (2nd ed.), Addison-Wesley, Massachusetts.

Neave, H. R. (1973) On using the Box-Muller transformation with multiplicative congruential pseudo-random number generators. Appl. Statist. 22.

Niki, N. (1979) Multi-folding the normal distribution and mutual transformation between uniform and normal random variables. Ann. Statist. Math., 31, 1, A.

仁木（1980）工学的乱数発生．統数研い報２７巻１号．

仁木（1983）パーソナル・コンピュータのための物理乱数発生器．（投稿中）

表１．代表的な発生法により作られた一様乱数列の短い構成単位の統計的性質

表中の数字は、１６進一様乱数をそれぞれの方法で発生し、連続する 4,096個ごとの各値（０から１５まで）の出現回数Ｎ(i) (i=0, 1,…,15)から

$$X^2 = \sum_{i=0}^{15} \{ N(i) - 256 \}^2 / 256$$

により求めた「適合度値」30,000個の度数分布を示す。最右欄は自由度１５のカイ二乗分布から求めた期待度数である。

テストされた一様乱数発生法は、以下に掲げた４種である。

乗算合同法：乗数＝39,894,229、法＝２**32
ｍ－系列：シフトレジスタ長＝５２１、伏見、手塚（1981）
ＲＮＤ関数：ＰＣ８００１のＢＡＳＩＣの組込み関数
物理乱数：統計数理研究所の物理乱数発生装置、石田ら（1972）

項	適合度値	乗算合同法	ｍ－系列	ＲＮＤ関数	物理乱数	期待度数
┌	0--1	0	0	0	0	.0
1	1--2	1	0	0	0	.9
│	2--3	8	4	45	5	11.2
└	3--4	65	68	72	52	55.8
2	4--5	187	169	75	164	168.3
3	5--6	363	363	366	385	371.4
4	6--7	643	689	625	656	662.9
5	7--8	945	1041	735	1089	1016.0
6	8--9	1334	1432	1158	1414	1388.0
7	9--10	1765	1754	1421	1745	1733.3
8	10--11	2009	2002	1738	2014	2014.4
9	11--12	2240	2240	2015	2186	2207.0
10	12--13	2256	2328	2316	2334	2301.9
11	13--14	2323	2266	2174	2296	2303.1
12	14--15	2233	2169	2394	2279	2223.4
13	15--16	2034	2055	2137	1990	2081.0
14	16--17	1865	1865	1716	1926	1895.6
15	17--18	1701	1628	1982	1638	1686.0
16	18--19	1545	1512	1655	1478	1468.0
17	19--20	1336	1228	1532	1218	1254.0
18	20--21	1065	1069	886	1098	1053.1
19	21--22	836	829	950	832	870.8
20	22--23	695	681	715	675	709.9
21	23--24	568	555	837	583	571.4
22	24--25	456	433	643	455	454.5
23	25--26	347	363	351	345	357.6
24	26--27	302	300	335	266	278.6
25	27--28	256	231	306	202	215.0
26	28--29	155	151	129	176	164.5
27	29--30	114	132	158	126	124.9

（次ページに続く）

（前ページの続き）

項	適合度値	乗算合同法	m－系列	ＲＮＤ関数	物理乱数	期待度数
28	30--31	99	100	119	89	94.1
29	31--32	71	85	101	83	70.4
30	32--33	54	53	87	62	52.3
31	33--34	39	57	44	36	38.7
32	34--35	26	36	28	32	28.4
33	35--36	18	33	56	23	20.7
34	36--37	15	19	29	12	15.1
35	37--38	9	19	14	6	10.9
└	38--39	5	10	0	9	7.8
┌	39--40	5	6	28	6	5.6
│	40--41	3	7	14	3	4.0
│	41--42	3	2	0	7	2.9
│	42--43	2	6	14	2	2.0
│	43--44	1	2	0	2	1.4
36	44--45	1	1	0	0	1.0
│	45--46	0	0	0	0	.7
│	46--47	1	1	0	1	.5
│	47--48	0	1	0	0	.3
│	48--49	0	1	0	0	.2
└	49--	1	4	0	0	.5
適合度（注１）		39.6	55.8	982.5	32.2	(.0)
確率　（注２）		.273	.014	.000	.605	(1.000)

（注１）：この表を、最左欄の『項』に従って、３６のクラスの度数分布表に再編する。そして、各欄の度数分布の期待度数分布に対する『適合度』を計算する。この行は、そのカイ二乗値を示す。

（注２）：自由度３５のカイ二乗分布の上側確率をしめす。

図１．Ｂｏｘ－Ｍｕｌｌｅｒ法にみられる離散近似の影響（Ｎｅａｖｅ効果）

Ｎｅａｖｅ（1973）が採上げたのは乗算合同法の特殊ケースに関してであるが、そうでない場合にも、下図ほどではないとしても、同じ現象が生じる。

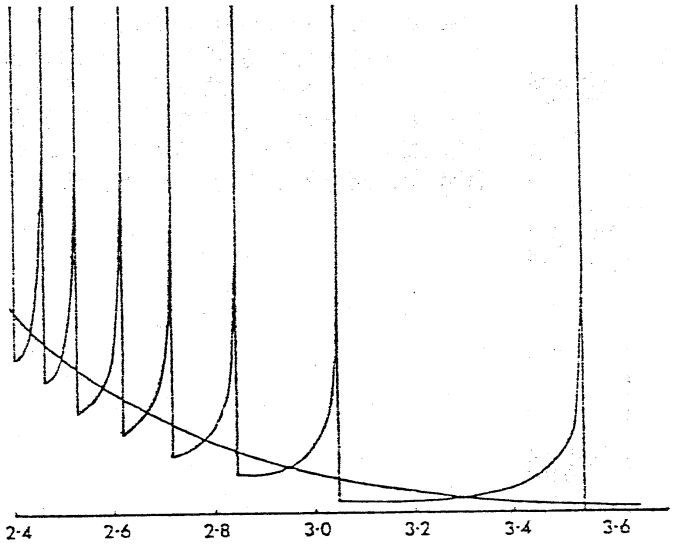


図２．逆関数法における離散近似の影響

正規分布の分布関数の逆関数Ｆ⁻が計算できれば、（0,1)一様乱数ｕから、

$$x = F^{-}(u)$$

により、正規乱数ｘが直ちに得られる。しかし、ｕが離散値をとる場合にはｘも離散値をとり、下図のように、確率密度を下げる代わりにｘのとり得る値の間隔を拡げているにすぎない。ここでは、ｕのとり得る値が .0001 刻みの場合を例示してある。

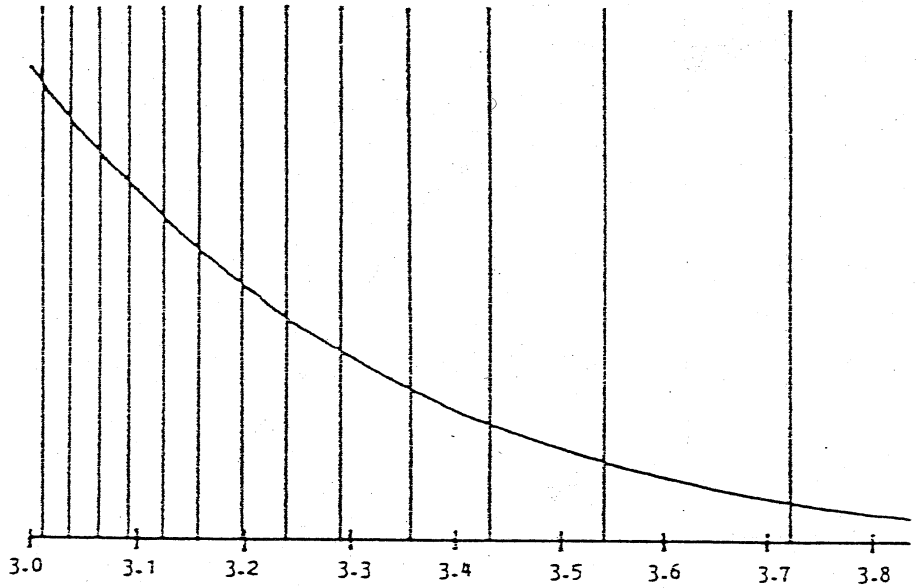
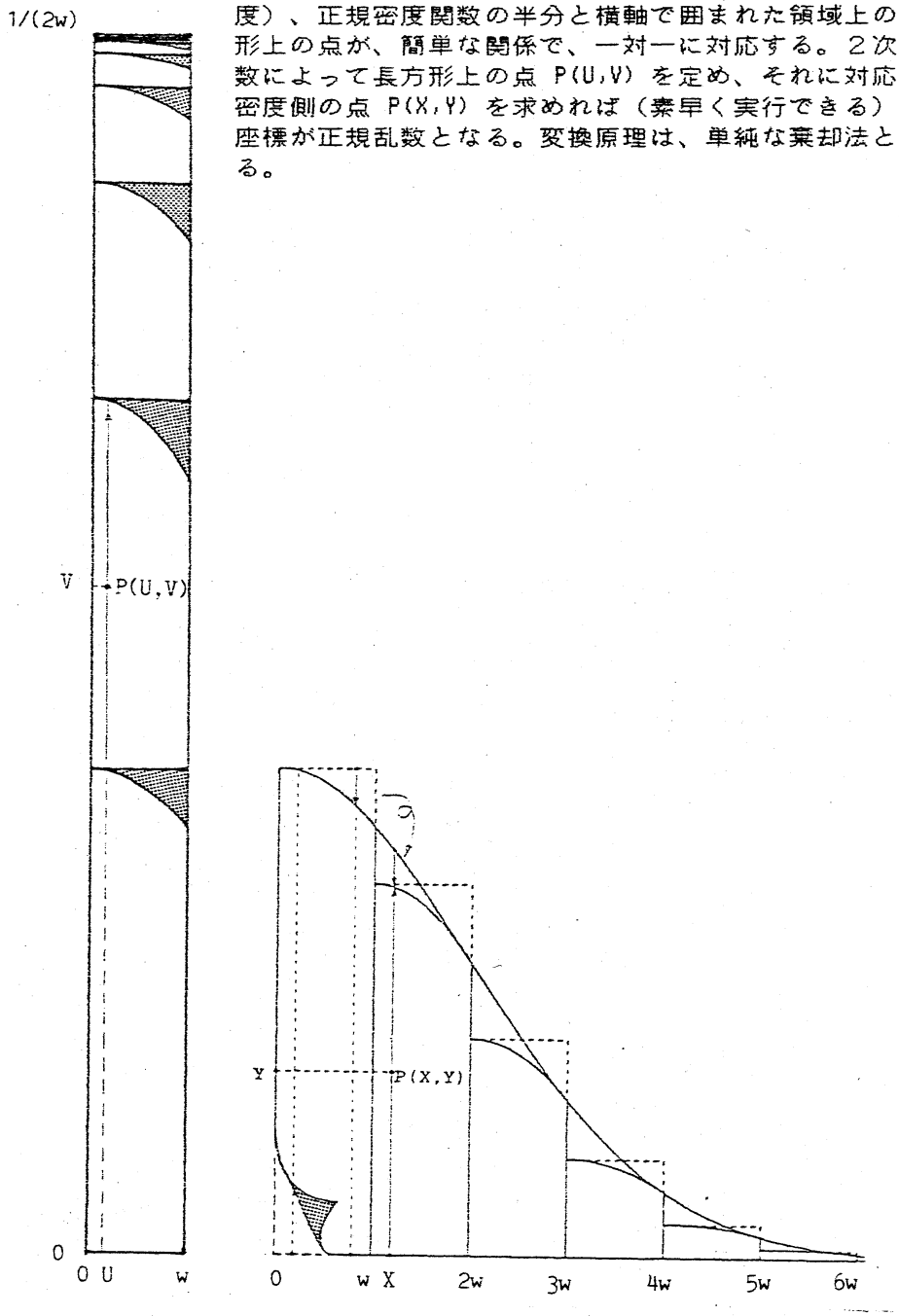


図３．正規乱数の発生法〔仁木（1979）〕



wの値をある程度小さく（例えば1/4 ）にとれば、実際上全く無視できる裾の部分を除いて（w＝1/4 のとき10の -34乗程度）、正規密度関数の半分と横軸で囲まれた領域上の点と長方形上の点が、簡単な関係で、一対一に対応する。２次元一様乱数によって長方形上の点 P(U,V) を定め、それに対応する正規密度側の点 P(X,Y) を求めれば（素早く実行できる）、そのｘ座標が正規乱数となる。変換原理は、単純な棄却法と同じである。