乱数の品質

樋口さぶろお

龍谷大学理工学部数理情報学科

計算科学☆実習 B　L11(2019-07-23 Tue)

最終更新: Time-stamp: "2019-07-23 Tue 18:11 JST hig"

今日の目標

- 乱数の品質のチェック方法を説明できる
- 適切な乱数生成ライブラリを選択できる

L10-Q1

Quiz 解答:逆関数法

$0 \leq r < 2, 0 \leq y < 1$ において,

$$\frac{3\sqrt{2}}{8}\sqrt{r}\mathrm{d}r = 1\mathrm{d}y$$
$$\frac{3\sqrt{2}}{8}\frac{1}{3/2}r^{3/2} = y + C$$

$y = 0$ のとき $r = 0$ より, $C = 0$ で, $g(y) = 2y^{2/3}$. このとき $y = 1$ のとき $r = 2$ も自動的に満たされる.

L10-Q2

Quiz 解答:逆変換法による擬似乱数生成

$$
\begin{aligned}
f_{\mathrm{R}}(r)\mathrm{d}r &= f_{\mathrm{Y}}(y)\mathrm{d}y \\
-\frac{200}{21}\frac{1}{r^3}\mathrm{d}r &= \mathrm{d}y \\
\frac{200}{21}\frac{1}{2}r^{-2} &= y + C
\end{aligned}
$$

$y = 0$ のとき $r = -2$ より, $C = \frac{4}{21}$. $r^2 = 100(4 + 21y)^{-1}$.
$r = \pm[100(4 + 21y)^{-1}]^{1/2}$.
$r(0) = -5, r(1) = -2$ に注意すると, $r = g(y) = -10(4 + 21y)^{-1/2}$
L10-Q3

Quiz 解答:逆関数法

$2 \leq r < 5, 0 \leq y < 1$ において,

$$\frac{1}{3}\mathrm{d}r = 1\mathrm{d}y$$

$$\frac{1}{3}r = y + C$$

$y = 0$ のとき $r = 2$ より, $C = 2$ で, $g(y) = 3y + 2$. このとき $y = 1$ のとき $r = 5$ も自動的に満たされる.

L10-Q4

Quiz 解答:一様分布と正規分布

$-a = b = \sqrt{3}$.

L10-Q5

Quiz 解答:棄却法による乱数生成

ソースコード 1: 棄却法

```
1  extern double getuniform();
2  double f(double x);
3
4  double getrandom(){
5    double x,y;
6    double a=1.0, b=2.0;
7    double M=2.0; /* = max f = f(2) */
8
9    while(1){
10     x=a+(b-a)*getuniform(); /* U(1,2) */
11     y=M*getuniform(); /* M(0,2) */
12     if(y<f(x)){
13       break;
14     }
15   }
16   return x;
17 }
18
19 double f(double x){
```

```
20    return 2*(x-1); /* assume 1<=x<2*/
21  }
```

(x, y) は $(1, 2) \times (0, 2)$ の面積 2 の長方形に一様に分布する．このうち，$f(x) > y$ の面積 1 の領域の (x, y) に対しては x が返される．よって，`double getuniform()` 4 回の呼び出しに平均して 1 回乱数が返される．

ここまで来たよ

9 略解:逆関数法・棄却法による連続型乱数の生成

10 乱数の品質
- 乱数の品質

double getuniform() を疑おう

これまで使っていた `double getuniform()` は本当に独立同分布 $\mathrm{U}(0,1)$ にしたがっているのだろうか? → No.

```
1 #include <stdlib.h>
2 double getuniform(){
3   return rand()/(RAND_MAX+1.0);
4 }
```

見るからにだめ

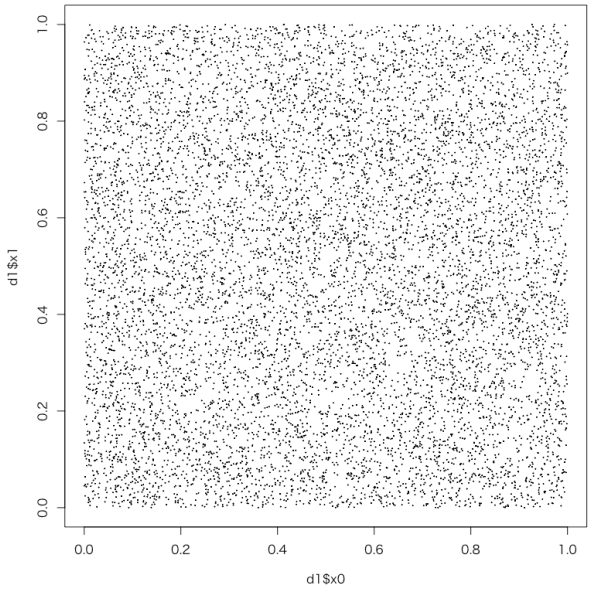
- 実数って言うけど, `1.0/(RAND_MAX+1.0)` 幅の格子点.
- 乱数表の長さだけ呼び出したら, その後は繰り返しになる.

疑うべき点

- ヒストグラムは $\mathrm{U}(0,1)$ に似てる? $P(\texttt{getuniform()} < y) = y$?
- $\mathrm{E}[\texttt{getuniform()}] = 0.5$? $\mathrm{V}[\texttt{getuniform()}] = 1/12.0$?
- `getuniform()` と次の `getuniform()` は独立同分布にしたがう?

擬似乱数列の独立性

$R(0)$ =getuniform() と $R(1)$ の散布図



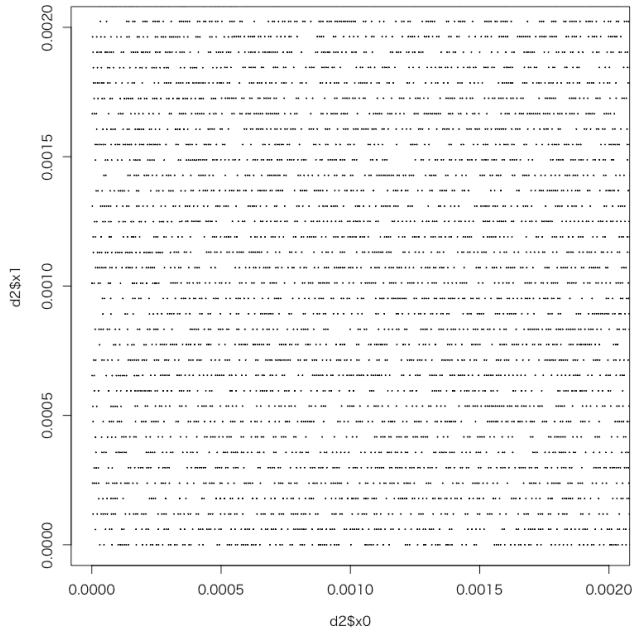
一様, 独立っぽい.

$N = 10000$ で 標本共分散 $S_{R(0)R(1)} = 0.0002553512.$

前園確率統計定理 4.3,5.3

しかし, 独立なら Cov=0 だけど, Cov=0 だから独立とは言えないのだった.

もっとサンプルサイズを大きくして, $[0, 0.02) \times [0, 0.002)$ 部分を拡大.



$R(0)$ と $R(1)$ は独立でない.

実は `int rand()` は乱数表を見ているわけではない.

```
1 srand(seed);
2 for(t=0;t<100;t++){
3    print("%d\n",rand());
4 }
```

は次と同じようなことをやってる.

線形合同法

```
1 int a=214013,b=2531011,m=32768; /* Visual C++*/
2 x=seed;
3 for(t=0;t<100;t++){
4    x = (a*x + b)%m; /* ax + b mod m */
5    printf("%d\n",x);
6 }
```

前の項から次の項が決まる. これ自身, 時系列
周期は最大でも $m \leq$ `RAND_MAX`.

整数論 (離散数学)

今後, 乱数を使うときのアドバイス

言い訳 `int rand()` は `stdlib.h` に入ってるから C では OS やコンパイラによらず使える.. そのため, この授業では `int rand()` から `double getuniform()` を作って使ってきた. しかし, `rand()` は低品質.

一般的なおすすめ

各処理系に一様分布や主要な確率分布にしたがう乱数生成関数が備わっている. 品質は様々.

- 近江崇宏先生 (東京大学) C 言語による乱数生成 `http://www.sat.t.u-tokyo.ac.jp/~omi/random_variables_generation.html`

MT=Mersenne Twister という数学的理論を使った高品質な乱数があれば/信頼できる 3rd party が提供していれば, それを使おう. 整数論 (離散数学)

- 松本真先生 (広島大学) Mersenne Twister `http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/mt.html`

予習復習問題のやり方+今後の予定

しばらく情報メディアセンターの Moodle で

`https://moodle.media.ryukoku.ac.jp`

Moodle App for iOS/Android

URL をきかれたら `https://moodle.media.ryukoku.ac.jp` で登録.