逆関数法による連続型擬似乱数生成

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 L09(2017-06-19 Mon)

最終更新: Time-stamp: "2017-06-19 Mon 13:00 JST hig"

今日の目標

- 与えられた確率密度関数 $f_R(r)$ にしたがう乱数を, $[0,1)$ 一様乱数から逆関数法で生成できる.
- 母ナントカと標本ナントカ, ラグランジュ表現とオイラー表現を対照して説明できる
- ペンギンの群れの配置をラグランジュ表現, オ

http://hig3.net

L08-Q1

Quiz(確率変数の変換)

あるクッキーマシンの作る正方形のクッキーの面積 (生地の量)Q は, 次の確率密度関数にしたがう (単位省略).

$$f_{\mathrm{Q}}(q) = \begin{cases} \frac{1}{36} & (64 \leq q < 100) \\ 0 & (\text{他}) \end{cases}$$

クッキーの一辺の長さは $R = g(Q) = \sqrt{Q}$ で与えられる (単位省略).

1. Q の母平均値と母分散を求めよう.
2. 確率 $P(Q > 82)$ を求めよう.
3. $f_{\mathrm{R}}(r)$ を求めよう.
4. R の母平均値と母分散を求めよう (2 つの方法で).
5. 確率 $P(R > 9)$ を求めよう (2 つの方法で).

Quiz 解答:確率変数の変換

1. $Q \sim \mathrm{U}(64, 100)$ より,

$$\mathrm{E}[Q] = \int_{-\infty}^{+\infty} q\ f_{\mathrm{Q}}(q)\ \mathrm{d}q = \tfrac{64+100}{2} = 82.$$

$$\mathrm{V}[Q] = \frac{1}{12}(100 - 64)^2 = 108.$$

2.

$$P(Q > 82) = \mathrm{E}[\mathbf{1}_{[Q>82]}(Q)] = \int_{82}^{100} \tfrac{1}{36}\ \mathrm{d}q = \tfrac{1}{2}.$$

一様分布で母平均値より大きい値を得る確率は $\frac{1}{2}$.

3. $f_{\mathrm{R}}(r)\mathrm{d}r = f_{\mathrm{Q}}(q)\mathrm{d}q$, $r = q^{1/2}$, $\frac{\mathrm{d}r}{\mathrm{d}q} = \frac{1}{2}q^{-1/2}$ より,

$$\begin{aligned} f_{\mathrm{R}}(r) =& \frac{1}{\frac{1}{2}\frac{1}{\sqrt{q}}} f_{\mathrm{Q}}(q) = 2r f_{\mathrm{Q}}(q) \\ =& \begin{cases} 2r \times \frac{1}{36} & (8 \leq r < 10) \\ 0 & (\text{他}) \end{cases} \end{aligned}$$

4. 母平均値 $\mathrm{E}[R]$ = 母期待値 $\mathrm{E}[g(Q)] = \mathrm{E}[\sqrt{Q}]$.

$$\begin{aligned} \mathrm{E}[R] =& \mathrm{E}[\sqrt{Q}] = \int_{-\infty}^{+\infty} \sqrt{q}\ f_{\mathrm{Q}}(q)\ \mathrm{d}q = \int_{64}^{100} \sqrt{q}\frac{1}{36}\ \mathrm{d}q = \frac{244}{27}. \\ \mathrm{E}[R^2] =& \mathrm{E}[Q] = 82. \\ \mathrm{V}[R] =& 82 - (\tfrac{244}{27})^2 = \frac{242}{729}. \end{aligned}$$

$f_{\mathrm{R}}(r)$ を先に求めたなら

$$\mathrm{E}[R] = \int_{-\infty}^{+\infty} r\ f_{\mathrm{R}}(r)\ \mathrm{d}r = \int_{8}^{10} r\ \tfrac{r}{18}\ \mathrm{d}r = \tfrac{244}{27}.$$

$$\mathrm{E}[R^2] = \int_{-\infty}^{+\infty} r^2\ f_{\mathrm{R}}(r)\ \mathrm{d}r = \cdots = 82.$$

5

$$P(R > 9) = P(Q > 81) = \tfrac{100-81}{36} = \tfrac{19}{36}.$$

$$P(R > 9) = \int_{9}^{10} \tfrac{r}{18}\ \mathrm{d}r = \tfrac{19}{36} > \frac{1}{2}.$$

L08-Q2
L08-Q3

Quiz 解答:確率変数の変換

$$f_Y(y) = \begin{cases} 1 & (0 \leq y < 1) \\ 0 & (\text{他}) \end{cases}$$ なので,

1.

$$\begin{aligned} \mathrm{E}[\mathrm{e}^{2Y}] &= \int_{-\infty}^{+\infty} \mathrm{e}^{2y} f_Y(y)\,\mathrm{d}y \\ &= \int_{-\infty}^{0} \mathrm{e}^{2y} \cdot 0\,\mathrm{d}y + \int_{0}^{1} \mathrm{e}^{2y} \cdot 1\,\mathrm{d}y + \int_{1}^{+\infty} \mathrm{e}^{2y} \cdot 0\,\mathrm{d}y \\ &= 0 + \frac{1}{2}(\mathrm{e}^2 - 1) + 0. \end{aligned}$$

2

$$P(R<2) = P(Y<\log 2) = \mathrm{E}[\mathbf{1}_{\log Y<2]}(Y)] = \int_{-\infty}^{\log 2} f_Y(y)\,\mathrm{d}y = \int_{-\infty}^{0} 0\,\mathrm{d}y + \int_{0}^{\log 2} 1\,\mathrm{d}y = 0 + \log 2$$

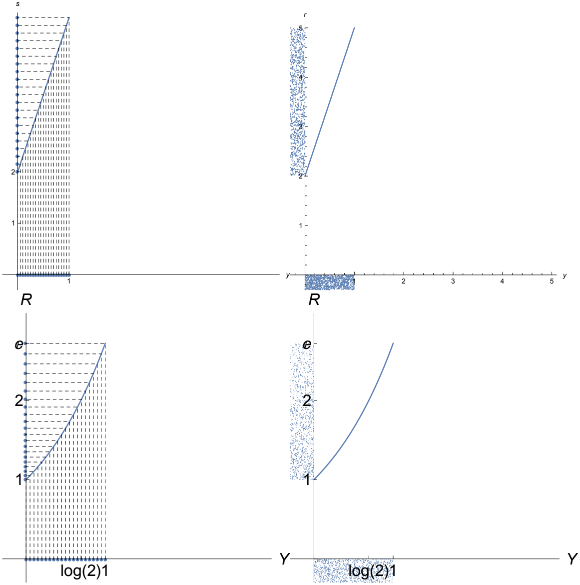
3 $f_R(r)\mathrm{d}r = f_Y(y)\mathrm{d}y$ より,

$$f_R(r) = \frac{1}{\frac{\mathrm{d}r}{\mathrm{d}y}} f_Y(y) = \mathrm{e}^{-y} \times f_Y(y) = \frac{1}{r} \times \begin{cases} 1 & (1 \le r < \mathrm{e}) \\ 0 & (他) \end{cases}$$

なお, この $f_R(r)$ を先に求めた場合は, 1,2 は次のように計算できる.

$$\mathrm{E}[R^2] = \int_{-\infty}^{\infty} r^2 f_R(r)\,\mathrm{d}r = \int_{1}^{\mathrm{e}} r^2 \cdot \frac{1}{r}\,\mathrm{d}r = \frac{1}{2}(\mathrm{e}^2 - 1).$$

$$P(R<2) = \int_{-\infty}^{2} f_R(r)\,\mathrm{d}r = 0 + \int_{1}^{2} \frac{1}{r}\,\mathrm{d}r = \log 2 - 0.$$



g の傾き大 $\Leftrightarrow$ $f_R(r)$ 小.

ここまで来たよ

9 連続型擬似乱数の生成と変換

10 逆関数法による連続型擬似乱数生成
- 逆関数法
- オイラー表現とラグランジュ表現

逆関数法

Q が一様分布にしたがうときでも，$R = g(Q)$ は一様分布にしたがうとはかぎらない．以下 $Q = Y$([0,1) 一様分布 $\mathrm{U}(0,1)$ で．

逆に，与えられた $f_{\mathrm{R}}(r)$ にしたがう擬似乱数を，うまい g で $R = g(Y)$ として作りたい!

変数変換 $r = g(y)$

$\mathrm{U}(0,1)$ にしたがう	y	0	$\to$	1
$f_{\mathrm{R}}(r)$ にしたがう	r	r_{min}	$\to$	r_{max}

$r_{\mathsf{min}}, r_{\mathsf{max}}$ は，$f(r) > 0$ となる r の下限，上限．

$r = g(y)$ はどんな関数?

$$
\begin{aligned}
\text{原理}\ f_{\mathrm{R}}(r)\,\mathrm{d}r =& f_{\mathrm{Y}}(y)\,\mathrm{d}y \\
\text{微分方程式}\ f_{\mathrm{R}}(r)\frac{\mathrm{d}r}{\mathrm{d}y} =& f_{\mathrm{Y}}(y)
\end{aligned}
$$

両辺を $y' = 0$ から y まで積分

$$
\begin{aligned}
\int_0^y f_{\mathrm{R}}(g(y'))\frac{\mathrm{d}r}{\mathrm{d}y}(y')\mathrm{d}y' =& \int_0^y f_{\mathrm{Y}}(y')\,\mathrm{d}y' \\
\int_{r_{\mathsf{min}}}^r f_{\mathrm{R}}(r')\mathrm{d}r' =& y
\end{aligned}
$$

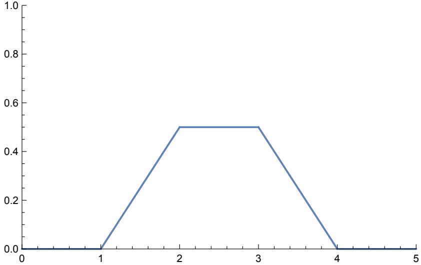
左辺を R の累積分布関数という．確率密度関数の原始関数.

累積分布関数

$$
F(r) = \int_{-\infty}^r f_{\mathrm{R}}(r')\,\mathrm{d}r' = P(R < r)
$$

累積分布関数 $F(r)$ の意味

r	$-\infty$	r_{min}		r_{max}	$+\infty$
$f_{\mathrm{R}}(r) = F'(r)$	$0 \leftarrow$	0	≥ 0	0	$\rightarrow 0$
$F(r)$					



逆関数法の手続き

さっきの式は,

$$F(g(y)) = y$$

乱数を作りたい確率変数 R の 累積分布関数 $F(r)$ の逆関数が $g(y)$.

逆関数法 (逆変換法)

$f_{\mathrm{R}}(r)$ に従う乱数を, $r = g(y)$ で $[0, 1)$ 一様乱数 Y から作るには, $g(y)$ を次の様に決めればいい.

1. R の累積分布関数 $F(r) = \displaystyle\int_{-\infty}^{r} f_{\mathrm{R}}(r')\,\mathrm{d}r'$ を計算する.
2. 範囲 $0 \le y < 1, r_{\mathrm{min}} \le r < r_{\mathrm{max}}$ で, $y = F(r)$ を解いて, 逆関数 $r = F^{-1}(y) = g(y)$ を求める.

動画 https://www.youtube.com/watch?v=cbgpdRW_5kQ

```
double getrandom(double y){ /*yは[0,1)一様乱数*/
  return g(y); /* 上で求めた式を書く */
}
```

L09-Q1

Quiz(逆変換法)

確率密度関数

$$f_R(r) = \begin{cases} \frac{1}{2}r & (0 \leq r < 2) \\ 0 & (\text{他}) \end{cases}$$

に従う乱数 R を, $[0,1)$ 一様乱数 y から $r = g(y)$ で作りたい. $g(y)$ を求めよう.

L09-Q2

Quiz($[a, b)$ 一様乱数の生成)

確率変数 $R \sim \mathrm{U}(-3, 2)$ を考える.

$$f(r) = \begin{cases} 1/5 & (-3 \le r < 2) \\ 0 & (他) \end{cases}$$

R に対応する疑似乱数を返す関数 `double getrandom(double y)` を書こう. ただし, y としては $[0, 1)$ 一様乱数を代入する.

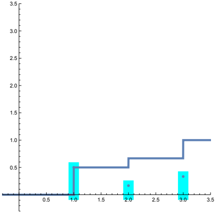
実は離散乱数の時も同じようなことやってた!

R	確率
1	$1/2$
2	$1/6$
3	$1/3$

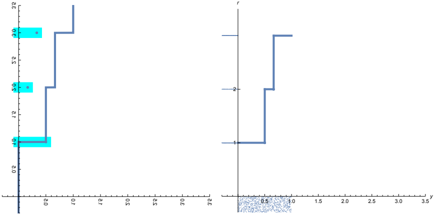
```
int getrandom(double y){
  int r;
  if(y<3/6.0){
    r=1;
  }else if(y<(3+1)/6.0){
    r=2;
  }else{
    r=3;
  }
  return r;
}
```

$$g(y) = \begin{cases} 1 & (0 \le y < 1/2) \\ 2 & (1/2 \le y < 2/3) \\ 3 & (2/3 \le y < 1) \end{cases}$$

確率関数と累積密度関数



累積密度関数の逆関数



ここまで来たよ

9 連続型擬似乱数の生成と変換

10 逆関数法による連続型擬似乱数生成

- 逆関数法
- オイラー表現とラグランジュ表現

実習課題の振り返り:2 つのタイプがあった!

- マルコフ連鎖の数値計算
  - markov, ...
  - 母ナントカ: 厳密. 確率の式を 1 回だけ計算. $p(x, t)$ は確率

    フォッカー-プランク, マスター方程式, 拡散方程式, 熱方程式
  - オイラー表現: 場所ごとに確率をカウント
- 確率シミュレーション
  - rw, sim, ...
  - 標本ナントカ: 標本サイズだけ乱数で実行を繰りかえして, 標本から推定. $X(t)$ は座標

    ランジュバン方程式

    ランダムウォーク
  - ラグランジュ表現:ウォーカーごとに座標をカウント

連続座標ランダムウォークの確率シミュレーション

```
int t;

int x;
int gentrandom(double y);

x+=getrandom(getuniform());
```

```
int t;

double x;
double gentrandom(double y);

x+=getrandom(getuniform());
```

とすれば離散座標と同じのりでできる.
連続座標ランダムウォークをマルコフ連鎖として扱うのは, 理論的には可能だけど, プログラムとしては書きにくい

なぜなら…

複数ウォーカーの確率シミュレーション (離散/連続座標)

同時に歩く KMAX=2 人. ウォーカーの座標: $X_1(t), X_2(t)$.

物理数学 I

```
#define KMAX 2   /* 人数 */
double x[KMAX];
double path[TMAX][KMAX];
for(n){ /*サンプル */
  t=0;
  for(k=0;k<KMAX;k++){ /*ウォーカー番号*/
    x[k]=初期位置;
    path[k][t]=x[k];
  }
  for(t){ /* 時間 */
    for(k=0;k<KMAX;k++){ /*ウォーカー番号*/
      x[k]=x[k]+乱数;
      path[k][t]=x[k];
    }
  }
}
double phi(double path[][KMAX], int tend, int kmax);
```

課題案

- 2 人のウォーカーが最接近するときの距離の母期待値は?
- 2 人のウォーカーが距離 1.0 以内で過ごす (通算) 時間の母期待値は?
- 2 人のウォーカーがすれちがう (=座標の大小が逆転する) 母比率は?

大注意: $X_k^{(i)}(t)$ $(i = 0, \ldots, N-1, k = 0, 1, \ldots, K-1, t = 0, 1, \ldots, T)$

- N は［　　　　］無関係に N 回のシミュレーションが繰りかえされる. N は試行の回数. n はレース番号=サンプル内データ番号.
- K は［　　　　］. k はウォーカー番号.
- T はランダムウォークの長さ, ［　　　　］. t は時刻.

L09-Q3

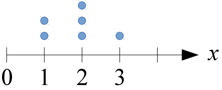
Quiz(2 人ウォーカーのサンプルパスの測定)

2 人ウォーカーの $0 \leq t \leq T = 9$ のランダムウォーク.
$(R_1(1), R_1(2), \dots, R_1(9)) = (+1, -1, -1, -1, -1, +1, +1, +1, -1)$,
$(R_2(1), R_2(2), \dots, R_2(9)) = (-1, -1, -1, +1, -1, -1, +1, -1, +1)$
というサイズ 1 の標本を考える. $X_1(0) = 0, X_2(0) = 10$ とする.

1. 2 人のウォーカーが最も近づいた時刻
2. 2 人のウォーカーが最も近づいたときの距離
3. 2 人の距離が 7 以下になった時間の長さ (=t の個数のこと)

ラグランジュ表現

確率は忘れて, ウォーカーが大勢いる状況をラグランジュ表現しよう.



数式的

$x^{(m)}(t)$: ウォーカー番号 m 番の, 時刻 t の座標.
上の状況なら
$x^{(0)}(t) = 1, x^{(1)}(t) = 2, x^{(2)}(t) = 2, x^{(3)}(t) = 3, x^{(4)}(t) = 1, x^{(5)}(t) = 2.$

C 的

`x[m]` ウォーカー番号 m 番の座標 (時刻 t とともに, この変数を更新)

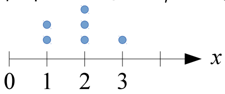
```
int x[6]; /*配列の宣言*/
```

または,

```
int x[]={1,2,2,3,1,2}; /*配列の宣言兼代入*/
```

オイラー表現

確率は忘れて, ウォーカーが大勢いる状況をオイラー表現しよう.



数式的

$P(x,t)$: 時刻 t に, 座標 x にいるウォーカーの人数.
上の状況なら
$P(0,t)=0, P(1,t)=2, P(2,t)=3, P(3,t)=1, P(\text{他},t)=0.$

C 的

`P[x]` 座標 x にいるウォーカーの人数 (時刻 t とともに更新)
`int P[100];` /*配列の宣言. $100-1=x$ 座標の上限*/
または
`int P[]={0,2,3,1,0,0,...};` /*配列の宣言兼代入*/
マルコフ連鎖の計算で使ってる `double p[]` は「いわば」 $p=P/N$, $N=6$ がウォーカーの合計人数.

L09-Q4

Quiz(ラグランジュ表現とオイラー表現)

(座標が整数値のみをとる離散型の) ランダムウォークを考える.
6 羽のペンギンが, 座標 $x = 0, 1, 2, \ldots, 9$ の範囲をランダムウォークする.
ある時刻 t に, $x = 1$ に 2 羽, $x = 3$ に 3 羽, $x = 8$ に 1 羽いるとする.

1. ラグランジュ表現を用いたとき, 配列 `x[]` のサイズはどれだけ必要か. また, 時刻 t に配列の各要素はどのような値をとるか.
2. オイラー表現を用いたとき, 配列 `P[]` のサイズはどれだけ必要か. また, 時刻 t に配列の各要素はどのような値をとるか.

配列のサイズとは, 元の型の変数を何個集めたかという個数. `int x[SIZE];` の `SIZE`.

ラグランジュ表現とオイラー表現によるプログラムの比較

	ラグランジュ表現	オイラー表現
空間	なんでも	有限個の場所
ウォーカーの区別	あり	なし
得意な問		
シューティング ブロック崩し テトリス ランダムウォーク		

L09-Q5

Quiz(オイラー表現とラグランジュ表現)

次のゲームのオブジェクトのうち, オイラー表現に適したもの (=ラグランジュ表現に適していないもの) を答えよう.

1. シューティングの自機
2. シューティングのミサイル
3. シューティングの雑魚キャラ
4. シューティングのラスボス
5. ブロック崩しのボール
6. ブロック崩しのラケット
7. ブロック崩しのブロック
8. テトリスの落下前のブロック
9. テトリスの落下後のブロック

L09-Q6

Quiz(ラグランジュ表現)

ランダムウォークのラグランジュ表現で, 時刻 t におけるウォーカーの座標 $X(t)$ の標本が配列 `x[SAMPLESIZE]` に格納されているとする.

```
#define SAMPLESIZE 6
double x[SAMPLESIZE];
```

1. 標本平均値 $\overline{X}$ を計算して `double ex;` に代入するプログラム (の一部) を書こう.
2. $X(t) \leq 5$ の標本比率を計算して `double px;` に代入するプログラム (の一部) を書こう.

両者を同時に計算する 1 個のプログラム (の一部) でもよい.

L09-Q7

Quiz(オイラー表現)

ランダムウォークのオイラー表現，または，マルコフ連鎖の数値解法のプログラムで，時刻 t において ウォーカーの座標が $X(t) = x$ である確率 $p(x,t)$ が，すでに計算され，配列 `p[x]` に格納されているとする．ただし，$x = 0, 1, \ldots, 19$.

```
#define XMAX 20
double p[XMAX];
```

1. 母期待値 $\mathrm{E}[X(t)]$ を計算して `double ex;` に代入するプログラム (の一部) を書こう．
2. 母比率 $P(X(t) \leq 5)$ を計算して `double px;` に代入するプログラム (の一部) を書こう．

両者を同時に計算する 1 個のプログラム (の一部) でもよい．

お知らせ

- 2017-06-21 水 3 初夏のプチテスト (プログラミング実技)
- 今日 2016-06-19 月まで 課題 p103 プチテスト準備に役立つフィードバック解読

- 月昼 樋口オフィスアワー (1-502)
- チューター/Math ラウンジ 月火水木昼 1-614