連続型確率変数の擬似乱数・自己回帰モデル

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 L07(2019-06-06 Thu)

最終更新: Time-stamp: ”2019-06-07 Fri 08:31 JST hig”

今日の目標

- 連続型一様分布 $\mathrm{U}(a, b)$ の擬似乱数を生成できる.
- 連続座標の RW のプログラムが書ける.
- RW の座標の母平均値母分散が求められる.
- 自己回帰モデル AR(m) とランダムウォークの関係を説明できる.

L06-Q1

Quiz 解答:離散的なランダムウォークの確率の転置推移確率行列

1.

$$\begin{pmatrix} 1 & \frac{1}{7} & 0 & 0 & 0 \\ 0 & \frac{4}{7} & \frac{1}{7} & 0 & 0 \\ 0 & \frac{2}{7} & \frac{4}{7} & \frac{1}{7} & 0 \\ 0 & 0 & \frac{2}{7} & \frac{4}{7} & 0 \\ 0 & 0 & 0 & \frac{2}{7} & 1 \end{pmatrix}$$

2. 略

L06-Q2

L06-Q3

Quiz 解答:偏微分方程式の条件チェック

$$u(x,t) = \mathrm{e}^{-18t}\sin(3x).$$

ここまで来たよ

7 略解:ランダムウォークの境界条件・偏微分方程式の数値計算

8 連続型確率変数の擬似乱数・自己回帰モデル
- 1 次関数で連続型一様擬似乱数の生成
- (連続型) ランダムウォークの座標の母平均値と母分散
- 自己回帰モデル AR(m)

(復習) 離散型と連続型の確率変数

離散型:確率分布, 確率関数 前園確率統計 §2.1

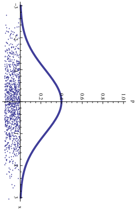
$f(r)$

得点 r	確率 $f(r)$
0	0.0667
1	0.2
2	0.3333
3	0.3
4	0.1

確率密度関数

連続型:確率密度関数 前園確率統計 §2.2

$f(r)$



- $f(r)$ が大きいほど, その値 r が
- $0 \leq f(r)$.
- $f(r)$ は 1 を超えることもある.

連続型確率変数の母期待値 (復習)

母期待値の定義 前園確率統計 §3.1

離散型確率変数　$\mathrm{E}[\phi(X)] = \sum_x f(x) \cdot \phi(x).$　$f(x)$:確率関数, 確率分布

連続型確率変数　$\mathrm{E}[\phi(X)] = \int_{-\infty}^{+\infty} f(x) \cdot \phi(x)\,\mathrm{d}x.$　$f(x)$:確率密度関数

母比率 (の一種)　$P(x_0 \leq X \leq x_1) = \mathrm{E}[\mathrm{I}_{[x_0 \leq X \leq x_1]}(X)] = \int_{x_0}^{x_1} f(x)\,\mathrm{d}x.$

全事象の確率　$1 = \mathrm{E}[1] = \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x.$

連続型一様分布 前園確率統計 §2.2(p.23) 確率統計☆演習 I(2018)L08

連続型一様分布 $\mathrm{U}(a,b)$

確率変数 X の確率密度関数が次で与えられるとき, X は区間 $[a,b)$ の連続型一様分布 $\mathrm{U}(a,b)$ に従うという.

$$f(x) = \begin{cases} C(\text{定数}) & (a \leq x < b) \\ 0 & (\text{他}) \end{cases}$$

L07-Q1

Quiz(一様分布)

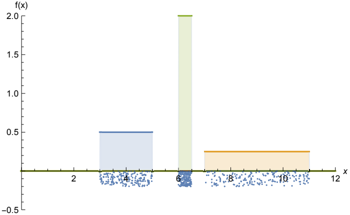
連続型確率変数 X が一様分布 $\mathrm{U}(c,d)$ にしたがう.

1. C を求めよう.
2. $\mathrm{E}[X]$ を求めよう.
3. $\sqrt{\mathrm{V}[X]}$ を求めよう.

$Y = aX + b$ の意味

X が 一様分布 $\mathrm{U}(r, s)$ にしたがうとき,
$Y = aX + b$ は 一様分布 $\mathrm{U}(ar + b, as + b)$ にしたがう.

$\mathrm{E}[aX + b] =$
$\sqrt{\mathrm{V}[aX + b]} =$

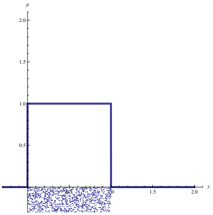


左から $X \sim \mathrm{U}(3, 5)$, $Z = \frac{1}{4}X + \frac{21}{4}$, $Y = 2X + 1$.

連続型確率変数に対応する擬似乱数

例 連続型一様分布 $\mathrm{U}(a,b)$ 前園確率統計 §2.2
$Y \sim \mathrm{U}(0,1)$ の確率密度関数

$$f(y) = \begin{cases} 1 & (0 \le y < 1) \\ 0 & (\text{それ以外}) \end{cases}$$



に対応する擬似乱数 ($[0,1)$ 一様乱数) は?
以後しばらく, Y と書いたら $Y \sim \mathrm{U}(0,1)$.

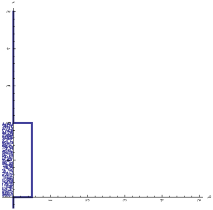
答: $\boxed{}$ そのもの 計算科学☆実習 B(2018)L01

$[0,2)$ 一様乱数を作るには?

$R \sim \mathrm{U}(a,b)$ の確率密度関数 $f_R(r)$ が与えられたとき, $R = g(Y)$ でそれにしたがう乱数を作ろう.

$$f(r) = \begin{cases} \boxed{?} & (0 \leq r < 2) \\ 0 & (\text{それ以外}) \end{cases}$$

```
double getrandom(double y){
  double r;
  r=???                    ;
  return r;
}
r=getrandom(getuniform());
```



y	r
0.31	0.62
0.82	1.64
0.49	0.98
0.04	0.08
0.60	1.20

$$r = g(y) = \boxed{???}$$

考え方 1: グラフ拡大縮小

$$f_{\mathrm{Y}}(y) = \begin{cases} 1 & (0 \leq y < 1) \\ 0 & (\text{他}) \end{cases}$$

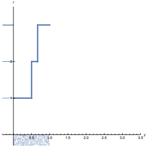
考え方 2: 母平均値や母分散や両端をあわせる 前園確率統計 §3.1,§3.2

離散型乱数の復習

今までは, Y を `int getrandom(double y)` で, 離散的な擬似乱数 R に'変換' していた.

R	確率
1	1/2
2	1/6
3	1/3

```c
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;
}
```

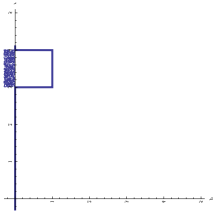


y	r
0.31	1
0.82	3
0.49	1
0.04	1
0.60	2

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

$[3,4)$ 一様乱数を作るには?

$$f(r) = \begin{cases} \boxed{?} & (3 \le r < 4) \\ 0 & (\text{それ以外}) \end{cases}$$



```
double getrandom(double y){
  double r;
  r=???          ;
  return r;
}
r=getrandom(getuniform());
```

y	r
0.31	3.31
0.82	3.82
0.49	3.49
0.04	3.04
0.60	3.60

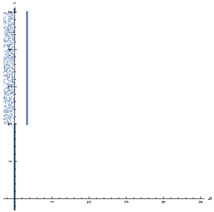
$r = g(y) = \boxed{???}$.

考え方 1: グラフ平行移動

$$f_{\mathrm{Y}}(y) = \begin{cases} 1 & (0 \le y < 1) \\ 0 & (\text{他}) \end{cases}$$

考え方 2: 母平均値や両端ををあわせる

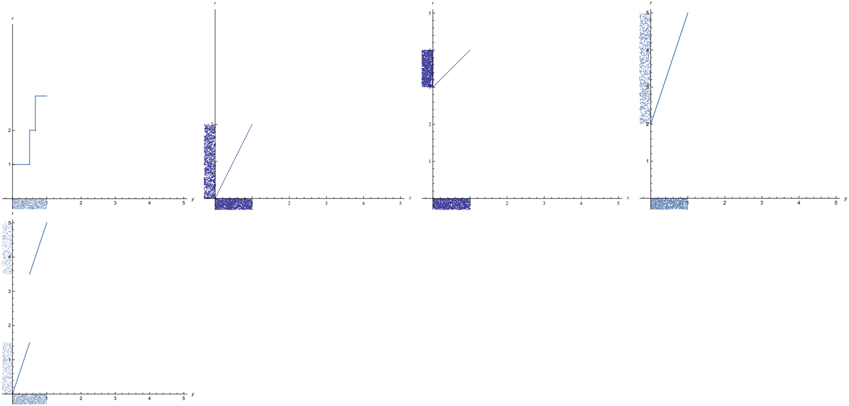
$[2,5)$ 一様乱数を作るには?



$$f(r) = \begin{cases} \frac{1}{3} & (2 \leq r < 5) \\ 0 & (\text{それ以外}) \end{cases}$$

$r = g(y) =???$

$g(y)$ の設計方法の解釈



自分の言葉でどうぞ

ここまで来たよ

(連続型) ランダムウォーク I

(連続型) ランダムウォークの定義

ランダムウォークの座標 $X(t)$: 次で決まる確率変数. $t = a, a+1, \ldots$
$R(t)$: 独立同分布に従う (連続型) 確率変数. $t = a+1, a+2, \ldots$

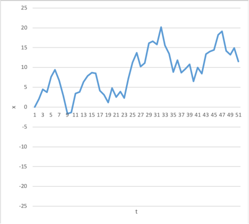
$$X(t) = X(t-1) + R(t), \qquad \text{初期条件 } P(X(a) = b) = 1$$

日本語で言うと,

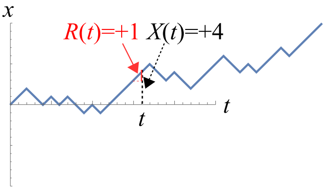
- x 軸上を移動するランダムウォーカーを考える
- ウォーカーは, 時刻 $t = a$ に, $x = b$ から出発する (確率が 1 である)
- ウォーカーは各時刻に, 確率変数 $R(t)$ だけ移動する

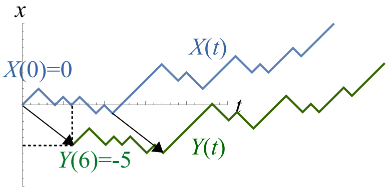
連続型 $R(t)$ の例 $R(t) \sim \mathrm{U}(c, d)$,
$\mathrm{E}[R(t)] = \mu = \frac{c+d}{2}$, $\mathrm{V}[R(t)] = \sigma^2 = \frac{(d-c)^2}{12}$

$X(T)$：時刻 T のランダムウォーカーの座標 (を確率変数とみたもの)
$(X(0), X(1), X(2), \ldots, X(T))$：パス (path) (を確率変数とみたもの)
連続座標ランダムウォーク



離散座標ランダムウォーク, 初期条件 $X(0) = 0$　初期条件 $X(6) = -5$





ランダムウォークの座標の母平均値と母分散 確率統計☆演習 I(2018)L06

$$
\begin{aligned}
X(t) =& X(t-1) + R(t) \\
=& (X(t-2) + R(t-1)) + R(t) \\
=& \cdots = X(a) + R(a+1) + \cdots + R(t)
\end{aligned}
$$

連続/離散型ランダムウォークの母平均値/母分散

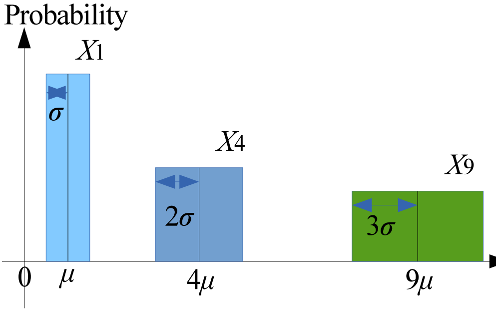
$\mathrm{E}[R(t)] = \mu, \mathrm{V}[R(t)] = \sigma^2$ のとき,

$$
\begin{aligned}
\mathrm{E}[X(T)] =& \mathrm{E}[X(a) + R(a+1) + \cdots + R(T)] = b + (T-a)\mu \quad (\mathrm{E}[] \text{ の性質}), \\
\mathrm{V}[X(T)] =& \mathrm{V}[X(a) + R(a+1) + \cdots + R(T)] \\
=& \mathrm{V}[R(a+1) + \cdots + R(T)] \quad (X(a) = b \text{ は定数}) \\
=& (T-a)\sigma^2 \quad (R(t) \text{ は互いに独立})
\end{aligned}
$$

$a = b = 0$ という簡単な場合, $X(t)$ の母平均値, 母分散は t に正比例する.

$X(t)$ の母標準偏差 $= \sqrt{\mathrm{V}[X(T)]} = \boxed{}$

ってことは，ランダムウォークの座標の確率分布の時間変化はこんな感じ?



L07-Q2

Quiz(ランダムウォーカーの到達点の座標の母平均・母分散)

時刻 $t=3$ に, $x=5$ から出発するランダムウォークの座標 $X(t)$ を考える.
時刻ごとに, 時刻 1 あたりの座標の増分 $R(t)$ は

$$\mathrm{E}[R(t)]=-3, \mathrm{V}[R(t)]=5$$

を満たす確率変数である. 時刻ごとの増分は独立である.

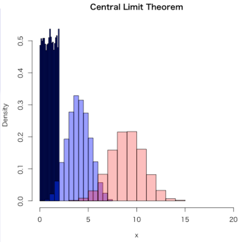
1. $X(20)$ の母平均値を求めよう.
2. $X(20)$ の母分散を求めよう.
3. $X(20)$ と同じ母平均値と母分散を持つ正規分布の確率密度関数 $f(x)$ とグラフを書こう.
4. $X(20)\geq -40$ の母比率を正規分布の上側確率 $Q(u)$ を使って近似的に表そう.

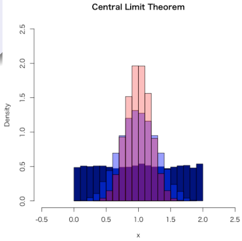
中心極限定理 前園確率統計 §3.4 確率統計☆演習 I(2018)L10

中心極限定理 (いいかげんバージョン) 前園確率統計 §3.4

$X_1, \ldots, X_n$ が母平均値 μ, 母分散 σ^2 の独立同分布に従うとき, $n \to +\infty$ で

- $U_n = X_1 + \cdots + X_n$, の確率分布は, の $\boxed{}$ に似る





ランダムウォークの場合.
$n \to t$, $X_i \to R(i)$. $U_n \to X(n)$
と思って適用できる.
問: さっきのランダムウォークの $X(20)$ の確率分布を描いて.

ここまで来たよ

7. 略解:ランダムウォークの境界条件・偏微分方程式の数値計算

8. 連続型確率変数の擬似乱数・自己回帰モデル
    - 1 次関数で連続型一様擬似乱数の生成
    - (連続型) ランダムウォークの座標の母平均値と母分散
    - 自己回帰モデル AR(m)

m 次の自己回帰モデル=AR モデル Auto Regression

m 次の自己回帰モデル AR(m)

$X(t)$: 連続型確率変数, $t = 0, 1, 2, 3, \dots$

$$X(t) = \sum_{k=1}^{m} a_k X(t-k) + R(t)$$

ただし, $R(t)$ は同分布にしたがい,

$$\mathrm{E}[R(t)] = 0, \qquad \text{WN1}$$

$$\mathrm{E}[R(t)R(s)] = \sigma^2 \times \delta_{t,s} = \sigma^2 \times \begin{cases} 1 & (t = s) \\ 0 & (\text{他}) \end{cases}, \qquad \text{WN2}$$

$$\mathrm{E}[R(t)X(s)] = 0 \quad (t > s). \qquad \text{PI}$$

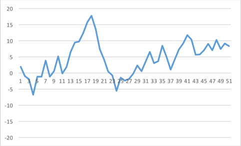
WN1, WN2 を満たす $R(t)$ をホワイトノイズ, 白色雑音 という.
PI はジャンプ幅が場所によらない (よって壁はない) という条件.

自己回帰モデルで記述できそうな現象の例

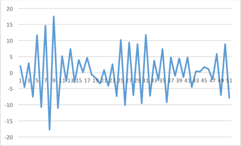
- $t =$ 日, $X(t) =$ 気温の, 平年平均気温との差, $0 < \phi < 1$.
- $t =$ シーズン, $X(t) =$ タイガースの勝率の 5 割との差, $-1 < \phi < 0$.

AR(1)

$\phi = 0.9,\ \sigma = 1$



$\phi = -0.9,\ \sigma = 1$



AR(1) モデルとランダムウォーク

AR(1) $a_1 = \phi$ (一般の値).

```
for(t){/*AR(1)*/
  x=phi*x+getrandom(getuniform());
}
```

AR(1) で $\phi = 1$ で, $R(t)$ が独立とすると, ランダムウォークで $\mathrm{E}[R] = 0$ で, 壁がないものを実現できる.

```
for(t){/*ランダムウォーク*/
  x=x+getrandom(getuniform());
}
```

$\mathrm{E}[R] = 0$ で壁のないランダムウォーク $\subset$ AR(1) $\subset$ 自己回帰モデル $\subset$ 時系列

L07-Q3

Quiz(AR(1) モデルの例)

AR(1) モデルで, 出発点を $X(0) = 100.0$ とする. 得られた乱数の値を, $R(1) = 15.0, R(2) = -8.0$ とする. $X(1), X(2)$ を, 3 つの場合 $\phi = 1.0, 0.9, -0.9$ について求めよう.

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

お知らせ

Moodle App for iOS/Android

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

- Math ラウンジ 1-538
- 樋口オフィスアワー火 5(1-507/1-542)
- 初夏のプチテスト (プログラミング実技) 2019-06-20 木 5
  - マルコフ連鎖
  - R + RStudio
  - 連続型ランダムウォークにおける母比率, 母期待値の推定