ランダムウォークのサンプルパスの標本抽出

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　L02(2021-04-15 Thu)

最終更新: Time-stamp: "2021-06-28 Mon 16:27 JST hig"

今日の目標

- ランダムウォークの座標を C で計算し, サンプルパスを Excel で描ける
- 標本から座標の母平均値を推定できる

L01-Q1

Quiz 解答:連続的な確率変数の母平均値・母分散・母標準偏差・確率 (一様分布)

1. $\mathrm{E}[\cos(\pi X)] = \int_{5/2}^{3} 2\cos(\pi x)\,\mathrm{d}x = [\frac{2}{\pi}\sin(\pi x)]_{5/2}^{3} = -\frac{2}{\pi}$
2. $P(\frac{22}{8} < X < \frac{23}{8}) = \mathrm{E}[\mathrm{I}_{[\frac{22}{8}<X<\frac{23}{8}]}(X)] = \int_{22/8}^{23/8} 2\,\mathrm{d}x = \frac{1}{4}$.

L01-Q2

Quiz 解答:擬似乱数の使いかた

ソースコード 1: 乱数

```
double getrandom(double y){
  if(y<0.3){
    return 0.4;
  }
  return 0.6;
}
```

L01-Q3

Quiz 解答:離散的な乱数の生成

```
int getrandom(double y){
  if(y<2.0/8.0){
    return 1;
  } else if (y<(2.0+1.0)/8.0){
    return 2;
  } else {
    return 3;
  }
}
```

L01-Q4
Quiz 解答:期待値

```
int getrandom(double y){
  if(y<2/13.0){
    return 0;
  } else if (y<(2+4)/13.0){
    return 3;
  } else {
    return 4;
  }
}
```

値 $R = 3$ が返される確率の検算.
$P(R = 3) = P(Y < 2/13$ でない かつ $Y < 6/13) = P(2/13 \le Y < 6/13) = \int_{2/13}^{6/13} 1 \,\mathrm{d}y = 4/13$.

ここまで来たよ

2 ランダムウォークと離散型擬似乱数

2 ランダムウォークのサンプルパスの標本抽出
- ランダムウォークを記述する言葉
- ランダムウォークのサンプルパスの抽出
- 確率シミュレーションによる推定

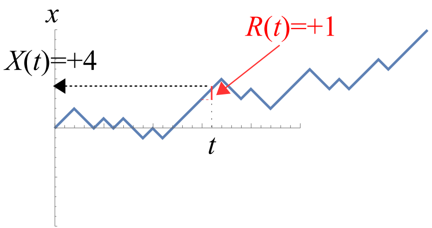
ランダムウォーク (確率過程の例)

確率変数 $X(t)$ がランダムウォークの座標 (数列 $\{X(t)\}$), 確率変数 $R(t)$ が独立同分布にしたがう (階差数列 $\{R(t)\}$)

$$\text{漸化式 } X(t) = X(t-1) + R(t) \quad (t = 1, 2, \ldots), \quad \text{初項 } X(0) = a.$$

現象の数理 A

サンプルパス (標本道)



例.

$R(t)$	確率
$+1$	p
-1	$q(=1-p)$

岩薩林 確率・統計ベルヌーイ分布, $n = 1$ 二項分布 (p.66)

どうやってコンピュータで?

近いうちにこんな問題

L02-Q1

Quiz(確率シミュレーション)

$t=6$ に $x=-5$ から出発したランダムウォーカーが, $t=20$ で, 領域 $x<0$ にいる確率を推定して出力するプログラムを書こう. ただし,

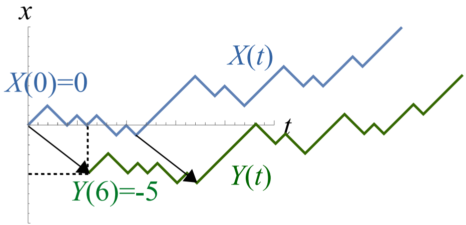
$$X(t)=X(t-1)+R(t)$$

で, `int getrandom(getuniform())` が独立同分布にしたがう確率変数 $R(t)$ を返すものとする. main と phi の中だけ書こう.

ランダムウォークの言葉づかいの習慣

$X(2)$：初期条件，ランダムウォーカーの出発点 (を確率変数とみたもの)

- 「ランダムウォーカーが時刻 $t = 6$ に $x = -5$ から出発した」⇔ $\boxed{P(X(6) = -5) = 1}$
- 「$t = 20$ で $x < 0$ にいる確率」⇔ $\boxed{P(X(20) < 0)}$



サンプルパス

ここまで来たよ

2 ランダムウォークと離散型擬似乱数

2 ランダムウォークのサンプルパスの標本抽出

- ランダムウォークを記述する言葉
- ランダムウォークのサンプルパスの抽出
- 確率シミュレーションによる推定

確率変数 R の標本

$R^{(n)}$

(n): サンプル内通し番号

n	
1	$R^{(1)}$, 改行
2	$R^{(2)}$, 改行
$\vdots$	$\vdots$
N	$R^{(N)}$, 改行

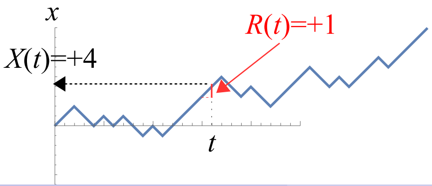
ランダムウォークの座標 $X(t)$ の標本

$X(t)^{(n)}$

t: 時刻, 数列の第 t 項

(n): サンプル内通し番号

	$t=0$	$t=1$	$\cdots$	$t=T$
$n=1$	$X(0)^{(1)}$,	$X(1)^{(1)}$,	$\cdots$	$X(T)^{(1)}$, 改行
$n=2$	$X(0)^{(2)}$,	$X(1)^{(2)}$,	$\cdots$	$X(T)^{(2)}$, 改行
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n=N$	$X(0)^{(N)}$,	$X(1)^{(N)}$,	$\cdots$	$X(T)^{(N)}$, 改行

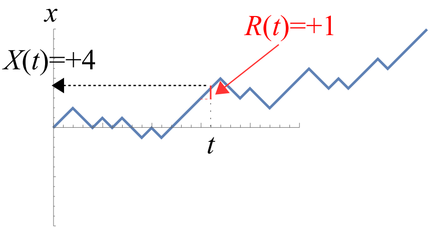


ランダムウォークの座標 $X(t)$ の標本 (標本サイズ 1)

まず, サンプルサイズ 1 で出力. $X(t)^{(1)}$

t: 時刻, 数列の第 t 項

	$t=0$	$t=1$	$\cdots$	$t=T$
	$X(0)$,	$X(1)$,	$\cdots$	$X(T)$, 改行



サンプルパス

C 言語で数列の計算

数列 $\{X(t)\}$ を定義の例.

数値計算法

$$\text{漸化式 } X(t) = X(t-1) + R(t) \quad (t = 1, 2, \ldots), \quad \text{初項 } X(0) = a.$$

数列の計算と出力

```
int x,t;

t=0;
x=a;
printf("%d,",x); /* t=0を特別扱い */
for(t=1;t<=100;t++){
   x=x+R(t);  /* X(t) を求めた */ /* R(t)階差数列 */
   printf("%d,",x);
}
printf("\n");
```

`int R(int t){return 3;}` なら $X(t)$ は初項 $\boxed{a}$ 公差 $\boxed{3}$ の等差数列.

ランダムウォーク $X(0), X(1), \ldots, X(T)$ のサイズ 1 の標本抽出

初項 $X(0) = a$.

```


                                        /*2*/


for(t=1;t<=T;t++){

                                        /*3*/
   x=x+getrandom(getuniform());
                                        /*4*/

}
                                        /*5*/
```

問: `srand(seed)`,x=a,t=0,`printf("%d,",x)` はどこ?

ランダムウォーク $X(0), X(1), \ldots, X(T)$ のサイズ N の標本抽出

初項 $X(0) = a$.

```
                                        /*1*/
for (n=0;n<N; n++){
                                        /*2*/

   for ( t=1;t<=T; t++){
                                        /*3*/
      x=x+getrandom(getuniform());
                                        /*4*/
   }
                                        /*5*/
}
                                        /*6*/
```

解説動画

https://www.youtube.com/watch?v=H-sIaVJmeiA

問: `srand(seed)`,x=a,t=0,`printf("%d,",x)` はどこ?

問: 他に何がいる?

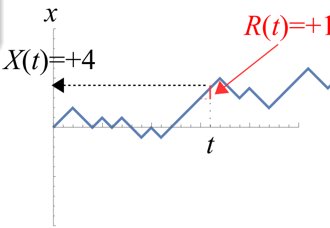
ここまで来たよ

ランダムウォークの座標の母期待値，母比率は?

$R(t)$ やランダムウォークの座標 $X(1000)$ は確率変数.

母ナントカの問

母ナントカを厳密に求めよう.



- 母平均値 $\mathrm{E}[R(t)]$, 母期待値 $\mathrm{E}[\mathrm{e}^{R(t)}]$, 母比率 $P(R(t) > 1)$
- 母平均値 $\mathrm{E}[X(1000)]$, 母期待値 $\mathrm{E}[\mathrm{e}^{X(1000)}]$, 母比率 $P(X(1000) > 1)$
- 母比率 $P(X(50) = 12$ かつ $X(100) = 25)$

別のタイプの問 (推定)

$R(t)$ の確率分布 $f(x)$ の式を知らない (例:誰かが作った getrandom の中身不明) けど, Excel データ (標本) だけはあるとする. 母ナントカを推定したい.

確率シミュレーション

母ナントカのコンピュータによる計算方法 (いま使わない)

$\mathrm{E}[X] = \sum_{x=1}^{3} x f(x)$ のような母ナントカの公式をプログラムで計算する.

$f(x)$ が求まったら苦労しないよ

最終的な誤差= $\boxed{\text{数値計算誤差}}$

別のタイプの作戦:確率シミュレーション

確率的現象を, 擬似乱数を使ってそのままコンピュータ上で再現し (simulate), 繰り返して実行して標本抽出して, 母ナントカを推定すること.

- いつでも推定はできちゃう
- 要 **コンピュータ or 奴隷**

最終的な誤差= $\boxed{\boxed{\text{統計誤差}}}$ + $\boxed{\text{数値計算誤差}}$

母期待値・母平均値・母比率の推定

母期待値の推定 岩薩林 確率・統計 §7.1 確率統計☆演習 I(2020)L12

$$\text{標本期待値}\ \overline{\phi(X)} = \frac{1}{N}\sum_{n=1}^{N}\phi(X^{(n)})$$

が 母期待値 $\mathrm{E}[\phi(X)]$ の ‘よい’ 推定値になっている.

確率変数 $Y = \phi(X)$ の母平均値の推定と思おう.

母平均値の推定 岩薩林 確率・統計 §7.1 確率統計☆演習 I(2020)L12

$$\text{標本平均値}\ \overline{X} = \frac{1}{N}(X^{(1)} + \cdots + X^{(N)}) = \frac{1}{N}\sum_{n=1}^{N}X^{(n)}$$

が, 母平均値 $\mathrm{E}[X]$ の ‘よい’ 推定値になっている.

母比率の推定 岩薩林 確率・統計 §7.3 確率統計☆演習 I(2020)L12

X のサンプルのデータ N 個中 k 個が「条件 $\cdots$ を満たす」とき，

$$\text{標本比率}\hat{p} = \frac{k}{N} = \frac{1}{N}\sum_{n=1}^{N} \mathrm{I}_{[\dots]}(X^{(n)})$$

が母比率 $P(\text{「条件}\cdots\text{」}) = \mathrm{E}[\mathrm{I}_{[\dots]}(X)]$ の ‘よい’ 推定値になっている．

L02-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)

仕組みのよくわからないランダムウォークで標本抽出したところ, $X(3)^{(n)} \quad n = 1, 2, \cdots, 10$ が

$$3, 3, 3, 1, 1, 1, 1, 1, -1, -3$$

だった.

1. 母平均値 $\mathrm{E}[X(3)]$ を点推定しよう.
2. 母分散 $\mathrm{V}[X(3)]$ を点推定しよう.
3. 母期待値 $\mathrm{E}[X(3)^3]$ を点推定しよう.
4. 母比率 $P(X(3) > 1)$ を点推定しよう.