確率過程とランダムウォーク

樋口さぶろお

龍谷大学理工学部数理情報学科

計算科学☆演習 II L02(2015-04-17 Fri)

最終更新: Time-stamp: "2015-04-25 Sat 09:57 JST hig"

今日の目標

- 確率過程とは何か説明できる
- ランダムウォークの確率シミュレーションのプログラムを書ける

http://hig3.net

L01-S1
Quiz 解答:擬似乱数の使いかた

ソースコード 1: 乱数

```
int getrandom(double y){
  if( y<1.0/3.0 ){
    return -1;
  } else if( y<1.0/3.0+1.0/2.0 ){
    return 0;
  } else {
    return +1;
  }
}
```

いろんな誤り

誤り S

```
if( y<1.0/3.0 ){
  /* A */
} else if (y<1.0/2.0){
  /* B */
} else {
  /* C */
}
```

のような誤り. A は $0 \leq y < 1/3$, B は $1/3 \leq y < 1/2$ のとき起きる. 確率は, この区間の長さに等しいので, A は $1/3$, B は $1/6$ の確率でおきることになっちゃう. この問題でいちばん考えてほしいポイントでした. もう 1 回考えてみてね.

誤り A

関数内で `y` が現れるたびに `getuniform()` がよばれると思っている誤り. 呼び出し元の関数 (例えば main) でいちどだけ `getuniform()` が実行され, その返り値の double の値が仮引数 `y` として渡されます.
`getrandom(getuniform())` で `getuniform()` が実行されるのは一度だけで, `getrandom(double y)` 内で y の値は変化しません.

誤り I

C では, $a < x < b$ みたいな不等式は書けません (正確には別の意味になっちゃう). 面倒でも&&で a<x && x<b と書かなきゃ.

誤り C

1/3 って書いてるけど, それ C では int 同士の整数の演算で 0 になっちゃうよ.

誤り F

$1/3 = 0.33$ と短い桁の小数で書いちゃった. 不正確すぎ.

誤り R

```
if( y<1.0/3.0 ){
  /* 何か */
} else if (y>=1.0/3.0 && y<5.0/6.0){
  /* 何か別 */
} else {
  /* また何か別 */
}
```

の y>=1.0/3.0 って無駄じゃない? ここに来るなら必ず成立してるよね.

ここまで来たよ

確率過程とは?

X:確率変数　確率統計 I

$\downarrow$

$\boxed{}$確率変数 $X(t)$ を確率過程という.

$(t = \ldots, -2, -1, 0, +1, +2, \ldots)$

$X(0), X(1), X(2), \ldots$ の間に '関係がある' 場合がおもしろい

関係がある $=$ 独立でない　確率統計 II

相関係数を求めたり, 回帰分析したりする X,Y みたいなもの.

$X = X(0), Y = X(1), \ldots$　確率統計 I

- 確率 $=$ probability
- 確率的な意味でランダムな $=$ stochastic
- 確率過程 $=$ stochastic process
- 確率微分方程式 $=$ stochastic differential equation 例. ファイナンスのブラック-ショールズ方程式

ここまで来たよ

ランダムウォーク

漸化式

$$X(t+1) = X(t) + R(t+1), \quad X(0) = 0.$$

これを解くと,

$$X(T) = 0 + \sum_{t=1}^{T} R(t).$$

ここで, $R(t)$ $(t = 1, 2, \ldots, T)$ は独立同分布に従う確率変数.
ということは, $X(t)$ は確率変数. ランダムウォークは確率過程の例.

- 独立: $R(t_1), R(t_2)$ が無関係
- 同分布 $R(t)$ の確率分布 $f(r)$ が共通

ランダムウォークの言葉づかいの習慣

$X(0)$：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 $t = 2$ に $x = 3$ から出発した」
$\Leftrightarrow x(2) = 3$
$\Leftrightarrow$

$X(T)$：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)
$x(T)$：時刻 T の到達点の座標 (の標本のデータ 1 個)
$(x(0), x(1), x(2), \ldots, x(T))$：サンプルパス (sample path)

ここまで来たよ

ランダムウォークのこんな問題?

$R(t)$	確率
-1	$q = 1 - p = \frac{2}{3}$
$+1$	$p = \frac{1}{3}$

座標 $X(t)$ について, 以下を厳密に求めよう. または推定しよう.

- $\mathrm{E}[X(2)]$, $\mathrm{E}[\mathrm{e}^{X(2)}]$, $X(2) > 1$ となる確率
- $\mathrm{E}[X(102)]$, $\mathrm{E}[\mathrm{e}^{X(102)}]$, $X(102) > 51$ となる確率
- $X(101) = 51$ かつ $X(102) = 50$ となる確率

厳密 $\rightarrow$ $\boxed{}$ 方法 1

推定 $\rightarrow$ $\boxed{}$ 方法 2

(方法 1-1) 母分布から厳密に手計算

$X(1) = R(1)$ の母分布

$X(1)$	確率
-1	$q = 1 - p = \frac{2}{3}$
$+1$	$p = \frac{1}{3}$

$X(2) = R(1) + R(2)$ の母分布

$X(2)$	確率	
-2	$\frac{2}{3} \cdot \frac{2}{3}$	$(-1) + (-1)$
0	$\frac{2}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{2}{3}$	$(-1) + (+1)$ or $(+1) + (-1)$
$+2$	$\frac{1}{3} \cdot \frac{1}{3}$	$(+1) + (+1)$

L02-Q1

ランダムウォークの確率と座標の期待値

$p = \frac{1}{3}$ のとき,

1. $P(X(3) = x)$ を求めよう (x は整数).
2. $\mathrm{E}[X(3)]$ を求めよう.
3. $\mathrm{V}[X(3)]$ を求めよう.
4. $X(3) > 1$ となる確率を求めよう.

(復習) 確率も期待値でかける

条件「$X(T) > 1$」となる確率$=\mathrm{E}[\mathbf{1}_{[X \text{ は } 1 \text{ より大}]}(X(T))]$.

$$\mathbf{1}_{[\text{条件}]}(x) = \begin{cases} 1 & (x \text{ に対して条件が成立する}) \\ 0 & (x \text{ に対して条件が成立しない}) \end{cases}$$

(方法 1-1) の調子で, 手計算の得意な人ならいくらでも. $\rightsquigarrow$ $\mathrm{E}[X(102)]$ とかでも求められる.

いくつかの作戦

- (方法 1-1) 手計算で $P(X(t)=x)$ を求める.
- (方法 1-3) 2 項定理から $P(X(t)=x)$ をいっきに式で書いちゃう 計算科学 II
- (方法 1-3') ランダムウォークの性質を使って, $\mathrm{E}[X(t)], \mathrm{V}[X(t)]$ を簡単に求めちゃう. 計算科学 II, 確率統計 II
- (方法 1-2) ランダムウォークの性質と中心極限定理で, $P(X(t)=x)$ を $T \to \infty$ の極限で近似的に求める. 計算科学 II, 確率統計 II
- (方法 1-4) 母関数の方法でなんでも求めちゃう 計算科学 II, 確率統計 II
- (方法 2) 計算機と乱数で標本抽出と推定でやっちゃえ $\to$ 確率過程の確率シミュレーション 計算科学 II

ここまで来たよ

1. 略解:ランダムウォークと擬似乱数生成

2. 確率過程とランダムウォーク
   - 確率過程とは?
   - ランダムウォーク
   - 母ナントカと標本ナントカ
   - 確率シミュレーション

(方法 2) 確率シミュレーション

擬似乱数を使ってサイズ N の標本 $X(T)^{(1)}, X(T)^{(2)}, \ldots, X(T)^{(N)}$ を作って, 母平均値 $\mathrm{E}[X(T)]$ を標本平均値

$$\overline{X(T)} = \frac{1}{N} \sum_{n=1}^{N} X(T)^{(n)}$$

で推定すれば?

復習

確率統計 I

- 母平均値は標本平均値で推定できる
- 母分散は標本分散で推定できる

同様に,

母期待値の推定

母期待値 $\mathrm{E}[\phi(X(T))]$ は, 標本期待値

$$\overline{\phi(X(T))} = \frac{1}{N}\sum_{n=1}^{N}\phi(X(T)^{(n)})$$

で推定できる.

理由:$Y = \phi(X)$ を確率変数と思えば, 母平均値の推定と同じこと.

$X(T)$ をいきなり返す `int getrandomT(double y)` を書くのはたいへん…

時間発展 (=漸化式) を適用して 1 項ずつ計算しちゃえ

$\rightarrow$ $\boxed{}$

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), くり返し実行して標本抽出し, 何かの母期待値を推定すること.

- とりあえずなんでも計算 (ていうか $\boxed{}$) できちゃう
- 要 $\boxed{}$

欲しい出力

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

t: $\boxed{}$

(n): $\boxed{}$

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

$X(1), X(2), \ldots, X(T)$ の標本を抽出するプログラム

```
                                   /*1*/
for(n<=N){
                                   /*2*/
  for(t<=T){
                                   /*3*/
    x=x+getrandom(getuniform());
                                   /*4*/
  }
                                   /*5*/
}
                                   /*6*/
```

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

標本から推定

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

Excel の関数: average, var, if(条件,,), sum

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，
$n < N = 10$, $t = T = 3$ の縦の列 $X(T)^{(n)}$ が

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

だった．

1. $\mathrm{E}[X(3)]$ を推定しよう．
2. $\mathrm{V}[X(3)]$ を推定しよう．
3. $\mathrm{E}[X(3)^3]$ を推定しよう．
4. $X(3) > 1$ となる確率 (比率) を推定しよう．

予習問題

火 23:55 締切の予習問題 x1 RaMMoodle
`https://el.math.ryukoku.ac.jp/moodle/`

manaba 出席カード提出
`https://attend.ryukoku.ac.jp`