ランダムウォークの座標の標本抽出と推定

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　L03(2016-04-25 Mon)

最終更新: Time-stamp: "2016-04-25 Mon 17:26 JST hig"

今日の目標

- 標本から, ランダムウォークの座標の母平均値, 母分散, 母期待値, 母比率が点推定できる.
- ランダムウォークの座標の標本抽出するプログラムが書ける.

http://hig3.net

L02-Q1

Quiz 解答:ランダムウォークの確率と座標の期待値

1. $$P(X(3)=x)=\begin{cases}\binom{3}{0}p^0(1-p)^3 & (x=0)\\ \binom{3}{1}p^1(1-p)^2 & (x=1)\\ \binom{3}{2}p^2(1-p)^1 & (x=2)\\ \binom{3}{3}p^3(1-p)^0 & (x=3)\end{cases}$$
2. $\mathrm{E}[X(3)]=(1-p)^3 0+3p(1-p)^2 1+3p^2(1-p)^1 2+p^3 3=3p.$
3. $\mathrm{E}[X(3)^2]=(1-p)^3 0^2+3p(1-p)^2 1^2+3p^2(1-p)^1 2^2+p^3 3^2=6p^2+3p.$
   $\mathrm{V}[X(3)]=\mathrm{E}[X(3)^2]-\mathrm{E}[X(3)]^2=3p(1-p).$
4. $\mathrm{E}[\mathbf{1}_{[X>1]}(X(3))]=(1-p)^3 0+3p(1-p)^2 0+3p^2(1-p)^1 1+p^3 1.$

L02-Q3

Quiz 解答:ランダムウォークの到達点の座標の母平均値・母分散

1. $\mathrm{E}[R(t)]=p\cdot(+1)+q\cdot 0=p.$

2. $\mathrm{E}[(R(t))^2] = p \cdot (+1)^2 + q \cdot 0^2 = p.$
   $\mathrm{V}[R(t)] = \mathrm{E}[(R(t))^2] - \mathrm{E}[R(t)]^2 = p - p^2 = pq.$
3. $\sqrt{\mathrm{V}[R(t)]} = \sqrt{pq}.$
4. $\mathrm{E}[X(t)] = tp.$
5. $\mathrm{V}[X(t)] = tpq.$
6. $\sqrt{\mathrm{V}[X(t)]} = \sqrt{t} \cdot \sqrt{pq}.$

L02-Q4

Quiz 解答:ランダムウォークと中心極限定理

1. $\mathrm{E}[X(20)] = 20 \cdot \mathrm{E}[R(t)] = -5.$ $\mathrm{V}[X(20)] = 20 \cdot \mathrm{V}[R(t)] = 2^2.$
2. $X(20)$ は独立同分布の和なので, $Z = \frac{X(20)-(-5)}{2}$ は標準正規分布 $\mathrm{N}(0, 1^2)$ に近似的に従う.
   $P(X(20) > -1) = P(Z > 2) = 1 - F(2) = 0.0228.$
3. $P(|X(20)| > 1) = P(Z < 2) + P(Z > 3) = F(2) + (1 - F(3)) = (1 - 0.0228) + 0.0013 = 0.9785.$

ここまで来たよ

こんなこと考えたいんだった

ランダムウォークの座標 $X(t)$ について,
(前回) 手計算で以下の母ナントカを求めよう.
(今回) 標本から以下の母ナントカを推定しよう.

- $\mathrm{E}[X(2)]$, $\mathrm{E}[\mathrm{e}^{X(2)}]$, $X(2) > 1$ となる確率
- $\mathrm{E}[X(1002)]$, $\mathrm{E}[\mathrm{e}^{X(1002)}]$, $X(1002) > 51$ となる確率
- $X(50) = 12$ かつ $X(100) = 25$ となる確率

方法 4:確率シミュレーション

母ナントカの公式は忘れたけどコンピュータはあるとする.
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?

点推定

標本平均値

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

が, 母平均値 $\mathrm{E}[X]$ の ‘よい’ 推定値になっている. 確率統計☆演習 I(2015)L10

(不偏) 標本分散

$$\text{(不偏) 標本分散 } S^2 = \frac{1}{N-1}[(X^{(1)} - \overline{X})^2 + \cdots + (X^{(n)} - \overline{X})^2]$$
$$= \frac{N}{N-1}\left[\frac{1}{N}\sum_{n}(X^{(n)})^2 - (\overline{X})^2\right]$$

が, 母分散 $\mathrm{V}[X]$ の ‘よい’ 推定値になっている. 確率統計☆演習 I(2015)L10

母期待値の推定

標本期待値

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

が 母期待値 $\mathrm{E}[\phi(X(T))]$ の ‘よい’ 推定値になっている.

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.

標本比率

サンプルのデータ N 個中 k 個が「…」であるとき,

$$\text{標本比率 } \hat{p} = \frac{k}{N}$$

が「…」の母比率 $\mathrm{E}[\mathbf{1}_{[\cdots]}(X)]$ のよい推定値になっている.

これらは点推定. 区間推定もあった母平均値 確率統計☆演習 I(2015)L11 母比率 確率統計☆演習 I(2015)L13 母分散 確率統計☆演習 I(2015)L13

L03-Q1

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，
$X(3)^{(n)}$ が

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

だった ($N = 10$).

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 (母比率) を推定しよう.

確率シミュレーション

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，母ナントカを推定すること．

- とりあえずなんでも計算 (ていうか　　　　) できちゃう
- 要

統計誤差 + 数値計算誤差

対立する方法

先週のように，$\mathrm{E}[X(t)]$ のような母ナントカを公式で書き，それをプログラムで計算する作戦．

数値計算誤差

ここまで来たよ

3 離散座標ランダムウォークの座標の確率・母平均値・母分散

4 ランダムウォークの座標の標本抽出と推定
- 標本からの推定
- 標本抽出するプログラム
- 母比率の推定の例

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

$X(0), X(1), X(2), \ldots, X(T)$ の標本を抽出するプログラム

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

  for(t=0;t<T;t++){
                                    /*3*/
    x=x+getrandom(getuniform());
                                    /*4*/
  }
                                    /*5*/
}
                                    /*6*/
```

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

問: 他に何がいる?

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, 真のときの式, 偽の時の式), sum

Excelを使わないで標本期待値の計算

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

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

sum1=0, sum1+=x*x*x ($\phi(x) = x^3$ のとき),
printf("%f",(double)sum1/N)?

ここまで来たよ

3 離散座標ランダムウォークの座標の確率・母平均値・母分散

4 ランダムウォークの座標の標本抽出と推定
- 標本からの推定
- 標本抽出するプログラム
- 母比率の推定の例

母比率の推定の例

L03-Q2

例題

$t = 2$ に $x = 10$ から出発したランダムウォーカーが, $t = 20$ で, 領域 $x < 0$ に到達する確率を推定しよう.
`double getuniform()` と (未知の)`int getrandom(double y)` は与えられているとして, `main()` だけをかけばよい.

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 $t = 2$ に $x = 3$ から出発した」 $\Leftrightarrow$ $\boxed{}$
- 「$t = 20$ で $x < 0$ に到達する」 $\Leftrightarrow$ $\boxed{}$

ある時刻の座標だけじゃなくて経路のサンプル

例題

$t=2$ に $x=10$ から出発したランダムウォーカーが, $0 \leq t \leq 20$ のいずれかの瞬間に, 領域 $x<0$ にいる確率

- これまで: 最終時刻の位置 $X(T)$ を集めたもの
- この例: サンプルパス $(X(0), X(1), \dots, X(t), \dots, X(T))$ を集めたもの

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

科目の 1 週間のタイムライン (修正)

1. 月昼 樋口オフィスアワー (1-502)
2. 月 15:20 予習復習問題 (e ラーニング) 解答~~1 回のみ~~ 解答何回でも. 最終解答で採点.
3. 月 4 講義 (7-002), quiz(参照あり)
4. このころ実習のタスク公開 目を通して考えておくこと推奨
5. 火 23:55 先週の課題の一部の提出締切
6. ~~水 13:35 予習復習問題 (e ラーニング) 解答何回でも~~
7. 水 3 実習 (1-609), quiz 返却
8. 水 23:55 今週の課題の一部の提出締切

実習室に行ったら, `http://hig3.net`→ 計算科学☆実習 B へ.

お知らせ

- 2016-05-11 水 3 実習の春のプチテスト
- チューター/Math ラウンジ 月火水木昼 1-614
- 統計検定

https://manaba.ryukoku.ac.jp
マイページの下の方に manaba 出席カード提出