ランダムウォークでモデルできる現象・サンプルパスの測定

樋口さぶろお

龍谷大学理工学部数理情報学科

計算科学☆演習 II L12(2015-07-03 Fri)

最終更新: Time-stamp: "2015-07-04 Sat 01:27 JST hig"

今日の目標

- ランダムウォークでモデル化できる現象の例を挙げられる
- サンプルパスが与えられたとき特徴量が計算できる

http://hig3.net

L11-Q2
Quiz 解答:母期待値の区間推定 (母分散未知)

1. 標本平均値 $\overline{Y} = -35$, 標本分散 $S^2 = 200$.
2. $$-35 - 2.776 \times \sqrt{\frac{200}{5}} < \mu < -35 + 2.776 \times \sqrt{\frac{200}{5}}$$
3. $$-35 - 4.604 \times \sqrt{\frac{200}{5}} < \mu < -35 + 4.604 \times \sqrt{\frac{200}{5}}$$

L11-Q3
Quiz 解答:母比率の区間推定

1. $\hat{p} = 0.2, \hat{p}(1 - \hat{p}) = 0.16.$

$$0.2 - 1.960 \times \sqrt{\frac{0.16}{1000}} < p < 0.2 + 1.960 \times \sqrt{\frac{0.16}{1000}}$$

2. $\hat{p} = 0.2, \hat{p}(1-\hat{p}) = 0.16.$

$$0.2 - 2.576 \times \sqrt{\frac{0.16}{100000}} < p < 0.2 + 2.576 \times \sqrt{\frac{0.16}{100000}}$$

L11-Q5

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. $X(20)$ は近似的に正規分布 $\mathrm{N}(-5, 2^2)$ に従うとみなせる. $Z \sim \mathrm{N}(0, 1^2)$ とすると, $X(20) = 2Z - 5.$ よって, 確率は, $P(X(20) > 0) = P(Z > \frac{0-(-5)}{2}) = 1 - F(2.5) = 0.0062.$
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

1. 略解:母期待値と母比率の区間推定・中心極限定理

2. ランダムウォークでモデルできる現象・サンプルパスの測定
   - ランダムウォークでモデルできる現象
   - サンプルパスの特徴量
   - もうかってる時間帯の長さ

現象の数理モデルとは?

自然や社会の現象を解析したいとき, 仕組みを, 些末な部分は無視して, 本質を数式やルールで表現して (モデリング), 本物はいったん忘れて数式やルール ((数理) モデル) だけを調べる作戦がある.

数理モデル

- 決定論的モデル
  - 微分方程式モデル (連続モデル)
    - ★ 常微分方程式モデル　(物理数学,) 数理モデル I,II, 計算科学 I,(現象の数学 B)
    - ★ 偏微分方程式モデル (離散モデル)　現象の数学 I, 現象の数学 B
  - 差分方程式モデル　多様体と力学系, 応用数理 A
- 確率論的モデル
  - 確率モデル (連続モデル, 離散モデル)　計算科学 II
  - 統計モデル (連続モデル, 離散モデル)　確率統計 I,II

ランダムウォークでモデルできる現象

- 破産ありの賭博
- クライマックスシリーズ
- 待ち行列の長さ
- 結晶成長 or 自動テトリスのゲームオーバー画面
- 株価変動
- ⋮

その他のモデリングの基礎にもなってる

クライマックスシリーズ

- チーム G とチーム T が 5 戦 (以下) して先に 3 勝した方が勝ち
- チーム G にアドバンテージ a 勝あり.

吸収壁ありランダムウォークとしてモデル化.
離散型確率変数

$$X(t) = (\text{チーム G の勝ち数}) - (\text{チーム T の勝ち数})$$

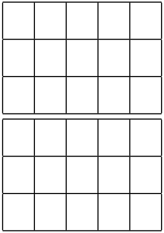
は次の法則にしたがう.

- 漸化式 $X(t+1) = X(t) + R(t+1)$
- 初期条件 $X(|a|) = a$
- 独立な離散型確率変数 $R(t)$

$$P(R(t) = r) = \begin{cases} p & (r = +1) \\ 1-p & (r = -1) \end{cases}$$

p:この対戦でのチーム G の勝率. とりあえず引き分け無視.

初期条件 $X(|a|) = a$
吸収壁: ここに来たらランダムウォーク終了.



- $X(2)$ は?
- G のシリーズ勝ち抜けの比率は?
- ウォークが終了する t の母平均値は? 母分散は?
- 最終戦まで行く確率は?
- 3 勝 1 敗で終わる確率は

簡単なら母ナントカを手計算で, 複雑なら確率シミュレーションで標本抽出して推定

L12-Q1

Quiz(吸収壁ありランダムウォークの確率の計算)

長さ $T = 4$ のランダムウォークで正の辺が $0, 1, 2, 3, 4$ 本となる確率を, 2^4 個のサンプルパスを描いて場合の数を数える (+節約) ことにより求めよう.

ここまで来たよ

1. 略解:母期待値と母比率の区間推定・中心極限定理

2. ランダムウォークでモデルできる現象・サンプルパスの測定
   - ランダムウォークでモデルできる現象
   - サンプルパスの特徴量
   - もうかってる時間帯の長さ

- 座標 $X(T)$ (T 固定)
  いままでは特定の $t = T$ を固定して, $X(T)$ の母ナントカを推定
- サンプルパス $(X(0), X(1), X(2), \ldots, X(T))$ という組

サンプルパスに関する量 (いくらでも作れる)

- $0 \leq t \leq T$ の範囲で, $X(t)$ の最大値
- $0 \leq t \leq T$ の範囲で, ($X(t)$ の最大値) $-$($X(t)$ の最小値)
- $0 \leq t \leq T$ の範囲で, もっとも多く訪れられた x

時刻 t や時間の長さ Δt も確率変数になる $\rightarrow$ 指数分布, 幾何分布 確率統計 II

- $0 < t \leq T$ の範囲で, 初めて $X(t) = 100$ となる t (定義になってる?)
- $0 < t < \infty$ の範囲で, 初めて $X(t) = 0$ に戻ってくる t (定義になってる?)
- $0 \leq t \leq T$ の範囲で, $X(t) > 3$ である t の個数

ここまで来たよ

1. 略解:母期待値と母比率の区間推定・中心極限定理
2. ランダムウォークでモデルできる現象・サンプルパスの測定
   - ランダムウォークでモデルできる現象
   - サンプルパスの特徴量
   - もうかってる時間帯の長さ

実験

$R(t) = (-1)^{\text{サイコロの目}}$ とすると

$$P(R(t) = r) = \begin{cases} \frac{1}{2} & (r = -1) \\ \frac{1}{2} & (r = +1) \\ 0 & (\text{他}) \end{cases}$$

となるはず (人力 getrandom).

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

で, $T = 9$ つまり $(X(0), X(1), \ldots, X(9))$ のサイズ 1 のサンプルを得よう

測定 1(復習)$X(T)$ の分布は?

これはサンプルパスでなく座標に対する問.

$\boxed{\text{送信方法}}$

$X(9)$	選択肢
-9	選択肢 0
-7	選択肢 1
-5	選択肢 2
-3	選択肢 3
-1	選択肢 4
$+1$	選択肢 5
$+3$	選択肢 6
$+5$	選択肢 7
$+7$	選択肢 8
$+9$	選択肢 9

測定 2:サンプルパスに対する問

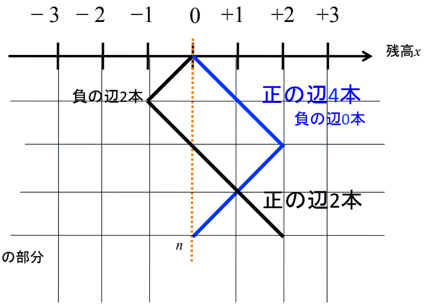
サンプルパスに対する問
$0 \leq t \leq T$ の範囲で $|X(t)|$ の最大値は?
送信方法 最大値は $u \to$ 選択肢 $u \quad (u = 1, 2, 3, 4, 5, 6, 7, 8, 9)$.

測定 3:正の辺負の辺の個数

サンプルパスに対する問
正の辺の個数:シリーズ

期間の長さ (だいたい)

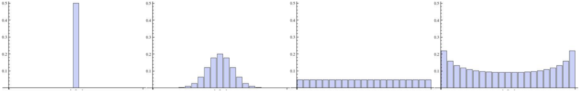
「期間n=4のうち黒字の期間がm」の確率
黒字の期間=勝ち越してる時間



正の辺
=赤線の右側の部分
=黒字の期間

どれになる?

$\boxed{\boxed{\text{送信方法}}}$ 選択肢 1,2,3,4
1,2,3,4



実際は?
$\boxed{\boxed{\text{送信方法}}}$
正の辺が u 本 $\to$ 選択肢 $u \quad (u = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9)$.

母ナントカのハイテクな計算によれば, $T \to +\infty$ で,

$$\text{確率密度関数} \quad f\left(\frac{\text{正の辺の個数}}{\text{辺の総数}\ T} = r\right) = \boxed{}$$

この確率密度関数の累積分布関数の関数形からついた名前: $\boxed{}$ 則

人生への教訓

L12-Q2

Quiz(与えられた正の辺の本数をもつサンプルパスの個数)

長さ $T = 4$ のランダムウォークで正の辺が $0, 1, 2, 3, 4$ 本となる確率を, 2^4 個のサンプルパスを描いて場合の数を数える (+節約) ことにより求めよう.

L12-Q3

Quiz(与えられた正の辺の本数をもつサンプルパスの個数)

長さ $T = 4$ のランダムウォークで正の辺が $0, 1, 2, 3, 4$ 本となる確率を, 2^4 個のサンプルパスを描いて場合の数を数える (+節約) ことにより求めよう.

お知らせ

Math ラウンジ=チューター月火水木昼.

スケジュール

2015-07-29 水 3 演習の真夏のプチテスト 35 ピーナッツ

2015-07-31 水 2 講義のファイナルトライアル (外部記憶ペーパー使用) 50 ピーナッツ

e ラーニング予習問題 次は 2015-06-30 火 23:55 締切

manaba 出席カード提出

`https://attend.ryukoku.ac.jp`