計算科学☆実習 B ファイナルトライアル (筆記)

樋口さぶろお[1] 配布: 2017-07-31 Mon 更新: Time-stamp: ”2017-08-08 Tue 06:59 JST hig”

ファイナルトライアル (筆記) 参加案内

1. 指定された用紙に解答しよう.
2. 過程も答えよう. 最終的な答えが正しいことがわかるような過程を記そう.
3. 問題文に現れない記号を使うときは, 定義を記そう.

1

過程不要

ランダムウォークの数式や計算機による計算では確率 p, 座標 x, 時間 t, 標本内のデータ番号 $n = 1, 2, \ldots, N$ (N は標本サイズ) が現れる.

p,x,t,n を使って下の問に答えよう.

1. ウォーカーの座標の確率を求めるための確率シミュレーションを行う. プログラム内の for 文のネスティングを,
   `for(x){ for(t){ for(p){}}}`
   のように答えよう (何重かも考えよう).
2. ウォーカーの座標の確率を求めるために, マルコフ連鎖の転置推移確率行列の積を数値計算する. for 文のネスティングを
   `for(x){ for(t){ for(p){}}}`
   のように答えよう (何重かも考えよう). ただし, 積の計算をする関数 `multiply_trans` の内部も含めて考える.
3. マルコフ連鎖に現れる分布 $\vec{p}(t) = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}$ の各成分 a_i は, p,x,t,n のどれに相当するか答えよう.
4. マルコフ連鎖に現れる分布 $\vec{p}(t) = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}$ の添字=成分を区別する番号 $i = 0, 1, 2, 3, 4$ は, p,x,t,n のどれに相当するか答えよう.
5. ランダムウォークに関係するある量 u は, ある極限で, 偏微分方程式

$$\frac{\partial u}{\partial t} = D \cdot \frac{\partial^2 u}{\partial x^2}$$

   にしたがう. この量 u は p,x,t,n のどれに相当するか答えよう.

[1]
hig@math.ryukoku.ac.jp, http://hig3.net(講義のページもここからたどれます), へや:1 号館 5 階 502

2

$\boxed{\textbf{過程不要}}$

次のうち, 正しいものに o, 正しくないものに x をつけよう.

1. 確率シミュレーションによる区間推定では, 標本サイズを大きくするほど, 信頼区間の長さは短くなる.
2. 推移確率行列の積を考えるマルコフ連鎖の数値計算では, 結果として母期待値や母比率の推定値である標本期待値や標本比率が求まる.
3. 複数ウォーカーの移動を記録するのにオイラー表現を使うと, $x = 0$ をちょうど 2 回訪れたウォーカーがいるかどうか, は判定できないことがある.
4. 複数ウォーカーの移動を記録するのにラグランジュ表現を使うと, 3 人のウォーカーが同じ時刻に同じ位置にいたことがあるかどうか, は判定できないことがある.
5. ランダムウォークをマルコフ連鎖として表現したとき, ある時刻 t の分布 $\vec{p}(t)$ は, オイラー表現に対応する.

3

$\boxed{\textbf{過程不要}}$

次のうち, 正しいものに o, 正しくないものに x をつけよう.

1. 時系列データの移動平均をとると, 一般的には, 元のデータよりもなめらかなデータが得られる.
2. 曜日の影響がある 1 日ごとの売上の時系列データで, 7 次の移動平均をとると, 曜日による周期的成分が打ち消されて, トレンドが見やすくなる.
3. 時系列データの移動平均をとると, トレンドが打ち消されて, ランダム効果が見やすくなる
4. 時系列データ $+1, -1, +1, -1, +1, -1, +1, \ldots$ の 1 次の自己相関係数は -1 である
5. 時系列データ $-1, +1, -1, +1, -1, +1, -1, \ldots$ の 2 次の自己相関係数は -1 である

4

$\boxed{\textbf{過程不要}}$

10 羽のペンギンが, 座標 $x = 0, 1, 2, \ldots, 9$ の範囲を (座標が整数値のみをとる) ランダムウォークをする.

ある時刻 t に, ペンギンが $x = 1$ に 2 羽, $x = 3$ に 3 羽, $x = 9$ に 5 羽いるとする.

1. ラグランジュ表現 (参考:実習の課題 contmrw02 のような表現) を用いたとき, 時刻 t に, 配列 `x[i]` の各要素はどのような値をとるか. 表で答えよう.
2. オイラー表現 (参考:実習の課題 markovpde01 のような表現) を用いたとき, 時刻 t に, 配列 `u[x]` の各要素はどのような値をとるか. 表で答えよう.

5

1. ある長さを表す離散型確率変数 R(cm) を考える. $R=r$ となる確率 $f(r)$ は

$$\begin{aligned} f(1) =& 0.4 \\ f(2) =& 0.3 \\ f(3) =& 0.2 \\ f(4) =& 0.1 \\ f(r) =& 0.0 \quad (r: \text{他}) \end{aligned}$$

   である. 母期待値 $\mathrm{E}[R]$, $\mathrm{E}[R^2]$, 母比率 $P(0 \leq R < 1.5)$ を求めよう.
2. ある長さを表す確率変数 R (cm) のサイズ 5 の標本がある. 標本中の i 番目のデータを $g(i)$ と書くと,

$$\begin{aligned} g(1) =& 0.4 \\ g(2) =& 0.3 \\ g(3) =& 0.2 \\ g(4) =& 0.1 \\ g(5) =& 0.0 \end{aligned}$$

   である. 標本平均値 $\overline{R}$, 不偏標本分散 S^2 , また, $0.0 \leq r < 0.15$ の標本比率 $\hat{p}$ を求めよう.

6

時刻 $t=100$ に 座標 $x=15$ から出発する, 座標が実数値をとるランダムウォークを考える. すなわち,

$$\begin{aligned} X(t+1) =& X(t) + R(t+1) \qquad (t \geq 100), \\ X(100) =& 15 \end{aligned}$$

である.

ここで, 連続型確率変数 $R(101), R(102), \ldots$ は独立同分布にしたがい, $\mathrm{E}[R(t)]=3, \mathrm{V}[R(t)]=4$ である.

1. $X(500)$ の母平均値を求めよう
2. $X(500)$ の母標準偏差を求めよう.
3. $500-100=400$ が十分に大きいと考えて中心極限定理を利用し, $X(500)>1200$ となる確率を近似的に求めて, 標準正規分布の累積分布関数 $\Phi(u)=\int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{z^2}{2}}\, \mathrm{d}z$ で表そう.

7

連続型確率変数 X は次の確率密度関数 $f(x)$ に従う.

$$f(x) = \begin{cases} \frac{1}{8}x & (0 \leq x < 4) \\ 0 & (\text{他}) \end{cases}$$

1. 母期待値 $\mathrm{E}[2X-3]$ を求めよう.
2. 母期待値 $\mathrm{E}[X^2]$ を求めよう.
3. 母分散 $\mathrm{V}[X]$ を求めよう.
4. 確率 $P(X<3)$ を求めよう.

8

超おんぼろな餃子の皮製造マシンは, 皮 1 枚あたり, X グラム (g) の原料を使う. X は, 4g 以上 9g 未満の実数値を同じ確からしさでとる (つまり X は連続型確率変数で, 一様分布 $\mathrm{U}(4,9)$ にしたがう). マシンは妙に正確で, 原料を決まった厚さの円板にのばすため, 円板の半径 R cm も確率変数であり,

$$X = (b\pi) \times R^2$$

という関係にある. ここで, b は餃子の皮の質量面密度で, いま $b = \frac{1}{2\pi}(\mathrm{g/cm^2})$ である.

1. 半径 R の母期待値を求めよう.
2. 半径 R が 5 cm 未満である確率 (母比率) を求めよう.
3. 半径 R の確率密度関数 $f(r)$ を求めよう.

9

次の確率密度関数 $f(r)$ にしたが連続型確率変数 R の乱数を生成したい.

$$f(r) = \begin{cases} \frac{1}{4} & (0 \leq r < 3) \\ \frac{1}{4} & (5 \leq r < 6) \\ 0 & (\text{他}) \end{cases}.$$

$[0,1)$ 一様乱数 Y に対して $R = g(Y)$ が $f(r)$ に従うような $g(y)$ を求めよう.

すなわち, 関数 `double getrandom(double y)` で return される値が $g(y)$ である.

10

次の確率密度関数 $f(r)$ にしたがう連続型確率変数 R の乱数を生成したい.

$$f(r) = \begin{cases} -\frac{1}{60}r^3 & (-4 \leq r < -2) \\ 0 & (\text{他}) \end{cases}.$$

逆関数法で $[0,1)$ 一様乱数 Y に対して $R = g(Y)$ が $f(r)$ に従うような $g(y)$ を求めよう.

すなわち, 関数 `double getrandom(double y)` で return される値が $g(y)$ である.

計算科学☆実習 B ファイナルトライアル (筆記) 略解

樋口さぶろお[2] 配布: 2017-07-31 Mon 更新: Time-stamp: ”2017-08-08 Tue 06:59 JST hig”

これは, 一部の過程のみ記した略解です. プチテストで, 受講者はすべての過程を記す必要があります.

配点 1-10:各 10 点, 計 100 点.

1

1. `for(n){ for(t){}}`
2. `for(t){ for(x){}}`
3. p
4. x
5. p

配点 1-5:各 2 点, 計 10 点.

講評 1,2 は設問方式が難しかった? 普通にプログラムしたときの for 文のネスティングと, そのときのカウンタ変数の意味を書いてねということです. これまでの例で, 確率が for 文のカウンタになってることなんてあった?

確率シミュレーションに標本 (n) がなかったり, 転置推移確率行列の積の計算に標本があったりしてはいけません.

2

o x o x o

配点 1-5:各 2 点, 計 10 点.

3

o o x o x

配点 1-5:各 2 点, 計 10 点.

講評 4,5 は次数を 1 ずれておぼえてた人には申し訳ない配点になってます.

hig@math.ryukoku.ac.jp, http://hig3.net(講義のページもここからたどれます), へや:1 号館 5 階 502.

4

1.

i	0	1	2	3	4	5	6	7	8	9
x	1	1	3	3	3	9	9	9	9	9

(順序は問わない. 要素 x の各値の個数があっていればよい)

2.

x	0	1	2	3	4	5	6	7	8	9
u	0	2	0	3	0	0	0	0	0	5

(順序までこの通りである必要がある)

配点　1,2:各 5 点, 計 10 点.

講評　簡単すぎましたね.

5

1. $\mathrm{E}[R] = 0.4 \times 1 + 0.3 \times 2 + 0.2 \times 3 + 0.1 \times 4 = 2(\mathrm{cm})$.
   $\mathrm{E}[R^2] = 0.4 \times 1^2 + 0.3 \times 2^2 + 0.2 \times 3^2 + 0.1 \times 4^2 = 5(\mathrm{cm}^2)$.
   $P(0 < R < 1.5) = P(R = 1) = 0.4$.
2. $\overline{R} = \frac{1}{5}[0.4 + 0.3 + 0.2 + 0.1 + 0.0] = 0.2(\mathrm{cm})$.
   $S^2 = \frac{1}{5-1}[(0.4-0.2)^2 + (0.3-0.2)^2 + (0.2-0.2)^2 + (0.1-0.2)^2 + (0.0-0.2)^2] = \frac{1}{4}(0.10) = 0.025(\mathrm{cm}^2)$.
   $r = \frac{1}{5}[0 + 0 + 0 + 1 + 1] = 0.4$.

配点　1 母平均値, 母期待値:各 2 点, 母比率 1 点, 2 標本平均値, 不偏標本分散:各 2 点, 標本比率 1 点, 計 10 点.

講評　今回は単位がなかったり間違ったりしていても減点していません.

6

1. $\mathrm{E}[X(500)] = \mathrm{E}[X(100)] + (500 - 100) \times \mathrm{E}[R(t)] = 15 + 400 \times 3 = 1215$.
2. $\mathrm{V}[X(500)] = (500 - 100) \cdot \mathrm{V}[R] = 1600 = 40^2$. よって, $\sqrt{\mathrm{V}[X(1200)]} = 40$.
3. 中心極限定理から, $Z = \frac{X(500)-1215}{40}$ は近似的に標準正規分布 $\mathrm{N}(0, 1^2)$ にしたがう. よって, $P(X(500) > 1200) = P(Z > \frac{1200-1215}{40}) = \int_{-\frac{3}{8}}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{z^2}{2}}\,\mathrm{d}z = 1 - \Phi(-\frac{3}{8}) = \Phi(\frac{3}{8})$.
   (別解) $P(X(500) > 1200) = \int_{1200}^{\infty} \frac{1}{\sqrt{2\pi \cdot 40^2}} \mathrm{e}^{-\frac{(x-1215)^2}{2 \cdot 40^2}}\,\mathrm{d}x = \int_{-\frac{3}{8}}^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{z^2}{2}}\,\mathrm{d}z$.

配点　1,2:各 3 点, 3(Z への変換),3(Φ による表示):各 2 点, 計 10 点.

講評　3 は毎回グラフに斜線を描いて意味を考えないと解けないくらい出題 (面積を求める部分) のバリエーションがあります.

7

1. $\mu = \mathrm{E}[X] = \displaystyle\int_{-\infty}^{+\infty} f(x)x\ \mathrm{d}x = \int_0^4 \frac{1}{8}x \times x\ \mathrm{d}x = \frac{8}{3}$. よって, $\mathrm{E}[2X-3] = 2\mathrm{E}[X] - 3\mathrm{E}[1] = 2\cdot\frac{8}{3} - 3\cdot 1 = \frac{7}{3}$.
   (別解) $\mathrm{E}[2X-3] = \displaystyle\int_{-\infty}^{+\infty} f(x)\times(2x-3)\ \mathrm{d}x$.
2. $\mathrm{E}[X^2] = \displaystyle\int_{-\infty}^{+\infty} f(x)\times x^2\ \mathrm{d}x = \int_0^4 \frac{1}{8}x\times x^2\ \mathrm{d}x = 8$.
3. $\mathrm{V}[X] = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = 8 - (\frac{8}{3})^2 = \frac{8}{9}$.
   (別解) $\mathrm{V}[X] = \mathrm{E}[(X-\mu)^2] = \displaystyle\int_{-\infty}^{+\infty} f(x)\times(x-\frac{8}{3})^2\ \mathrm{d}x$.
4. $P(X<3) = \mathrm{E}[\mathbf{1}_{[X<3]}(X)] = \displaystyle\int_{-\infty}^{+\infty} f(x)\times\mathbf{1}_{[X<3]}(x)\ \mathrm{d}x = \int_0^3 \frac{1}{8}x\ \mathrm{d}x = \frac{9}{16}$.
   (別解) グラフと面積で.

配点　1,3:各 3 点, 2,4:各 2 点, 計 10 点.

講評　確率統計☆演習 I(2016) のトライアルそのまま. おぼえてる〜? 1,3 は別解じゃないほうがはるかに楽.

8

1. $\mathrm{E}[R] = \mathrm{E}[\sqrt{X/(b\pi)}] = \sqrt{2}\mathrm{E}[\sqrt{X}] = \sqrt{2}\int_4^9 \frac{1}{5}x^{1/2}\ \mathrm{d}x = \frac{2\sqrt{2}}{15}(3^3-2^3) == \frac{38\sqrt{2}}{15} = 3.58$.
   (別解)3 の $f(r)$ を使って $\int_{\sqrt{8}}^{\sqrt{18}} \frac{1}{5}r\cdot r\ \mathrm{d}r$.
2. $P(4 \le X < 9) = 1$ なので, $P(R<5) = P(X < \frac{1}{2}\cdot 25^2) = 1$. (別解)3 の $f(r)$ を使って, $\sqrt{8} < \sqrt{18} < 5$ なので, $\int_{\sqrt{8}}^{\sqrt{18}} \mathbf{1}_{[R<5]}(r)\frac{1}{5}r\ \mathrm{d}r = \int_{\sqrt{8}}^{\sqrt{18}} \frac{1}{5}r\ \mathrm{d}r = 1$.
3. $4 \le x < 9$,$\sqrt{8} \le r < \sqrt{18}$ の範囲で, $\frac{1}{5}\ \mathrm{d}x = f(r)\ \mathrm{d}r$ で
   $\frac{\mathrm{d}x}{\mathrm{d}r} = r$ より,

$$f(r) = \begin{cases} \frac{1}{5}r & (\sqrt{8} \le r < \sqrt{18}) \\ 0 & (\text{他}) \end{cases}$$

配点　1:4 点, 2,3:各 3 点, 計 10 点.

講評　出題予告してたんだけど正解率は低かった, てことは授業内での説明が足りてなかったってことですね. Quiz L08-1 で, $Q = X$, $R = R$, $g(x) = \sqrt{2x}$ に相当する問題です.

2 で, 母比率が 1 より大きかったらどこかおかしいと思わなければいけません. 被積分関数が非零になるのは, $f(x)$ が正のところと, $\mathbf{1}_{[X<5]}(x)$ が 1 のところとの共通部分です.

$R \sim \mathrm{U}(2\sqrt{2}, 3\sqrt{2})$ や $\mathrm{E}[R] = f(\mathrm{E}[X])$ のようなまちがいは NG アンサーです.

9

$R \sim \mathrm{U}(0,4)$ なら $g(y) = 4y$ だが, 問題では, $r \geq 3$ 部分が 2 だけ平行移動されている.

$P(R < 3) = \frac{3}{4} = P(Y < \frac{3}{4})$ だから, $0 < y < \frac{3}{4}$ のとき $0 < r < 3$, $\frac{3}{4} < y < 1$ のとき $3 + 2 < r < 4 + 2$ となるようにすればよい.

よって,

$$g(y) = \begin{cases} 4y & (y < \frac{3}{4}) \\ 4y + 2 & (\frac{3}{4} < y) \end{cases}$$

配点　10 点.

講評　逆関数法を用いてもよい.

10

逆関数法による.

累積分布関数は $F(r) = \int_{-\infty}^{r} f(s)\,\mathrm{d}s = \int_{-4}^{r} -\frac{1}{60}\frac{1}{4}s^4\,\mathrm{d}s = -\frac{1}{240}(r^4 - 256)$.

$y = F(r)$ を r について解いて, 定義域 $0 \leq y < 1$ 値域 $-4 \leq r < -2$ の逆関数 $g(y)$ を求める.

$r^4 = 256 - 240y$.

値域に注意して符号を選ぶと, $g(y) = -(256 - 240y)^{1/4}$

配点　累積分布関数 $F(r)$:4 点, $g(y)$:6 点, 計 10 点.

講評　$g(y)$ が求まったら, $0 \leq y < 1$ で動かして R の値の範囲を $g(y)$ が動くかどうか検算しようよ〜

$(1+1)^{1/4} = 1^{1/4} + 1^{1/4}$ って成立しないよね〜

$g(y)$ の定義域は $[0, 1)$. それ以外の y に対して, 特に意味のある値の定め方はない. $g(y) = 0$ (他) ってしてる人いるけど, 確率密度関数じゃないんだから他はでないって意味じゃないよね? むしろ一瞬 $r = 0$ がとてもよくでそうな気分にならない?