

付録 A PC を用いた確率計算実習 [Mathematica]

Mathematica は、Wolfram 社の販売する数式処理ソフトです。グラフ化やシミュレーションまで、今後の学習・研究に役立つことでしょう。

- 情報科学部ではサイトライセンス契約で利用可能となっています。
- 利用方法の詳細は、次のページを参照してください。(講義ページからリンクされています)。
<http://www.oit.ac.jp/is/shinkai/lecture/mathematica.html>

演習室の PC からの起動方法

- Windows から、普通のソフトウェアと同じように、「プログラム」から「Wolfram Mathematica 8」を選択する。(下の方の「Mathematica カーネル」ではない。)
- 起動したら、5! と入力し、「shift+ 中央の enter」または「右端の enter」とキーボードを押して、120 と計算されるかどうか確かめよう。

A.1 課題

A.1 「徹底攻略 常微分方程式」(真貝著、共立出版、2010年8月)の p204–209 を熟読しながら、すべて実行して確かめよ*1。(p210–211 は不要)

A.2 (誕生日問題) N 人集まったとき、誕生日が同じ人がいる確率について考える。

- (1) 自分と同じ誕生日の人がいる確率を求め、 N を横軸としてグラフにせよ。確率が 0.5 を超えるのは、何人以上のときか。
- (2) N 人のなかで、同じ誕生日の人が少なくとも 1 組以上いる確率を求め、 N を横軸としてグラフにせよ。確率が 0.5 を超えるのは、何人以上のときか。

A.3 (サイコロをふる)

- (1) 上記ウェブページより、`dice.txt` を参照し、関数 `dice[n_, m_]` を Mathematica で定義せよ*2。例えば、5 回サイコロをふる試行を行うには、`dice[1,5]` を実行する。この試行を 10 回繰返すには `dice[10,5]` を実行すればよい。
- (2) 20 回サイコロを振り、1–6 の目について出現回数を記録せよ。
- (3) 上記ウェブページより、`dice1.txt` を参照し、関数 `dice1[n_]` を Mathematica で定義せよ。例えば、`dice1[100]` を実行すると、100 回サイコロをふったときの目の出た頻度が表示される。
- (4) サイコロをふって出る目の平均値が、試行回数を増やすと 3.5 に近づいていくことを示せ。

A.4 (2 項分布) 酔歩問題やサイコロ問題で登場する 2 項分布 $B(n, p)$ について。

- (1) 2 項分布 $B(n, p)$ で、 $X = k$ のときの確率を表示する関数を作成せよ。
- (2) 酔歩問題。左右どちらかに $1/2$ の確率で歩く酔っばらいが 10 歩後にいる位置の確率を求めよ。
- (3) (2) をグラフにせよ。

*1 「微積分学 I」「微分方程式」で行った練習と同じ。該当部分は(学内からなら)web ページで pdf が手に入る。

*2 平嶋先生作のスクリプト。

A.5 (標準正規分布)

- (1) 標準正規分布曲線 $f(z)$ をグラフで示せ.
- (2) $f(z)$ の $\alpha \leq z \leq 1000$ の部分の面積を計算する関数を作り, 値を標準正規分布表と比較せよ.
- (3) 知能指数が 125 以上の人は, 何 % いるだろうか.
- (4) 知能指数が 200 以上の人は, 何 % いるだろうか.

A.6 (Buffon の針) 平行線が $2h$ の間隔で無数に引かれている平面に, 長さ $2l$ (ただし $l < h$) の針を無作為に落とすとき, 針が平行線と交わる確率を p とする. p を用いると円周率 π が計算できる.

- (1) 上記ウェブページより, `buffon1.txt` を参照し, 関数 `buffon1[n_, length_]` を Mathematica で定義せよ. `buffon1[100, 0.7]` と実行せよ.
- (2) 講義で解説した例題 1.24 によれば, $\pi = 2l/ph$ である. シミュレーションを繰り返すと円周率 π が得られるだろうか. (`buffon1` では, $h = 1$ である).

A.2 方針とヒント

A.2 (誕生日問題)

- (1) 2 人なら, 相手が同じ誕生日の確率は, $\frac{1}{365}$. 3 人なら, $\frac{1}{365} + \frac{1}{365}$.
グラフをプロットするのは, `Plot[...]`
- (2) 余事象を考えることで, 確率は $1 - \frac{{}_{365}P_N}{365^N}$. これをプロットすればよい.
 ${}_{365}P_N$ を関数として組んでみよう.
`f[N]:=...`

A.3 (サイコロをふる)

- (1) 発生する乱数は毎回異なるので, 何回か試してみよう.
- (4) `dice1[]` のプログラムを使っても良いが, 次のコマンドを用いてグラフにすることもできる.
乱数を 1 から 6 までの範囲で, 10 個並べる命令文の例. (ついでにグラフも)

```
t1 = Table[Random[Integer, {1, 6}], {10}]
```

```
ListPlot[t1]
```

上記の Table に格納された数字の和を求めるためには,

```
Sum[t1[[i]], {i, 1, 10}]
```

付録 B PC を用いた統計計算実習 [Excel または Mathematica]

データ処理の 1 例として、「父親／母親の身長」と「息子／娘の身長」の相関を調べよう。データは、Excel または Mathematica 形式で、授業ページからリンクして置いてある。

B.1 課題

次のデータは、父親・母親とその成人した子供の身長データ (cm) である。2021-2012 年度の講義中に行ったアンケートから、兄妹あるいは姉弟の組み合わせの家族構成をもつデータ 197 家族分を抽出した。

- B.1** (1) 父親データの標本平均，標本分散，標準偏差を求めよ。
 (2) 母親データの標本平均，標本分散，標準偏差を求めよ。
 (3) 息子データの標本平均，標本分散，標準偏差を求めよ。
 (4) 娘データの標本平均，標本分散，標準偏差を求めよ。

B.2 父親・母親・息子・娘 のうちから 2 者を取り出し，相関係数をそれぞれ求めよ。

- B.3** (1) 父親・母親・息子・娘 のうちから 2 者を取り出し，回帰直線を求めよ。
 (2) 父親・母親の身長に対する息子の身長，および娘の身長について，重回帰分析をせよ。

x	y	z	w	x	y	z	w	x	y	z	w	x	y	z	w	x	y	z	w							
父	母	息子	娘	父	母	息子	娘	父	母	息子	娘	父	母	息子	娘	父	母	息子	娘							
1	165	156	168	154	41	171	164	178	159	81	164	166	172	163	121	175	160	175	158	161	173	164	173	165		
2	169	163	166	155	42	170	150	170	160	82	175	160	178	163	122	165	154	169	148	162	182	162	170	161	162	
3	173	159	174	159	43	164	160	180	163	83	155	145	159	150	123	166	160	161	152	163	168	150	166	150	163	
4	174	161	173	158	44	175	155	170	157	84	175	164	164	158	124	170	155	166	152	164	170	155	176	160	164	
5	175	158	172	162	45	175	158	173	175	85	173	163	171	165	125	180	147	160	155	165	167	156	171	157	165	
6	174	150	171	153	46	171	170	172	155	86	165	160	170	165	126	170	146	170	155	166	170	169	170	167	166	
7	175	155	160	155	47	155	160	177	154	87	160	150	166	153	127	173	160	175	170	167	175	158	180	160	167	
8	171	163	169	164	48	175	150	178	170	88	167	153	172	165	128	165	165	170	150	168	178	156	173	163	168	
9	175	158	170	155	49	171	165	160	158	89	175	154	171	150	129	173	162	186	163	169	165	158	161	153	169	
10	178	148	165	164	50	165	153	172	154	90	176	150	174	168	130	169	157	173	160	170	172	174	170	157	170	
11	163	163	161	152	51	175	156	174	164	91	178	163	173	165	131	174	162	178	160	171	172	155	170	160	160	
12	170	155	165	150	52	175	163	179	167	92	180	165	175	165	132	176	156	168	158	172	165	155	164	150	172	
13	175	167	171	161	53	168	160	173	166	93	167	155	170	158	133	170	160	170	158	173	169	161	172	162	173	
14	172	161	170	164	54	160	165	170	165	94	168	160	163	161	134	180	165	170	170	174	164	165	180	160	174	
15	174	161	176	158	55	165	150	167	155	95	165	159	160	155	135	173	154	172	156	175	171	163	181	153	175	
16	173	160	175	162	56	170	164	167	165	96	168	150	165	158	136	178	158	172	156	176	168	165	170	160	176	
17	175	165	175	165	57	166	153	163	153	97	167	165	176	160	137	170	165	177	160	177	159	157	165	160	177	
18	173	152	163	162	58	164	166	173	165	98	164	150	160	160	138	170	160	168	157	178	171	156	179	160	178	
19	169	152	167	153	59	161	153	165	149	99	170	164	170	167	139	164	160	171	161	179	160	150	165	145	179	
20	180	165	174	165	60	175	155	168	166	100	155	165	152	160	140	176	160	172	165	180	183	160	168	164	180	
21	172	150	162	153	61	175	152	168	157	101	174	160	175	162	141	175	160	170	163	181	173	163	170	152	181	
22	161	157	171	154	62	160	155	165	160	102	160	165	180	155	142	173	155	160	155	182	167	164	173	160	182	
23	168	150	170	153	63	163	154	173	160	103	174	153	170	156	143	178	164	186	166	183	178	159	170	160	183	
24	172	162	174	160	64	170	154	167	152	104	170	160	163	150	144	170	165	173	160	184	171	155	162	156	184	
25	178	155	172	168	65	175	165	173	160	105	171	158	172	153	145	175	160	174	161	185	160	167	168	153	185	
26	165	154	173	163	66	167	153	160	136	106	173	156	178	152	146	170	160	172	165	186	173	158	177	160	186	
27	163	158	172	160	67	170	165	169	156	107	170	150	165	150	147	170	150	170	160	187	175	160	178	160	187	
28	160	150	170	150	68	172	147	169	145	108	168	154	172	154	148	175	165	176	165	188	170	163	173	165	188	
29	177	156	172	156	69	172	168	160	153	109	165	155	176	160	149	169	158	165	155	189	170	150	165	155	189	
30	169	150	169	160	70	170	160	171	157	110	175	163	175	159	150	187	164	180	171	190	175	160	165	155	190	
31	175	155	168	160	71	174	170	185	165	111	175	150	170	159	151	160	154	166	158	191	172	155	175	160	191	
32	170	158	172	162	72	170	161	169	155	112	175	150	178	170	152	173	156	180	156	192	181	158	172	170	192	
33	167	158	171	162	73	168	158	168	158	113	165	150	166	160	153	175	163	165	140	193	167	158	166	160	193	
34	168	160	175	162	74	172	165	164	157	114	170	170	175	152	154	165	155	170	155	194	179	165	178	166	194	
35	170	155	164	152	75	169	162	172	165	115	160	162	171	152	155	170	162	172	160	195	165	155	168	160	195	
36	163	155	158	155	76	175	165	171	162	116	170	160	171	185	156	159	160	167	164	196	182	155	183	155	196	
37	170	165	178	170	77	169	162	180	163	117	179	170	182	165	157	171	156	166	154	197	167	166	177	155	197	
38	168	165	179	168	78	178	159	175	160	118	176	165	174	167	158	174	167	177	158							
39	171	164	177	170	79	189	170	179	169	119	176	160	180	160	159	175	160	162	164							
40	174	157	168	157	80	175	165	169	165	120	168	148	169	160	160	170	160	163	153							

B.2 ヒント

■平均, 積和

n 個のデータ (x_i, y_i) , $(i = 1, \dots, n)$ が与えられているとき, 平均と積和は, 次のように定義される.

$$\begin{aligned}\bar{x} &= \frac{1}{n} \sum_i x_i, & S_{xx} &= \sum_i (x_i - \bar{x})^2 \\ \bar{y} &= \frac{1}{n} \sum_i y_i, & S_{xy} &= \sum_i (x_i - \bar{x})(y_i - \bar{y})\end{aligned}$$

■相関係数

x と y の対からなる標本 (x_i, y_i) ($i = 1, \dots, n$) の相関係数 r は

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_i (x_i - \bar{x})^2 \right] \left[\sum_i (y_i - \bar{y})^2 \right]}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \quad (\text{付録 B.1})$$

■最小2乗法による回帰直線解析

n 個のデータ (x_i, y_i) , $(i = 1, \dots, n)$ が与えられているとき, これらのデータ分布を, もっとも良く近似する直線 (回帰直線) を

$$y(x) = ax + b \quad (\text{付録 B.2})$$

とすると,

$$a = \frac{S_{xy}}{S_{xx}}, \quad b = \bar{y} - a\bar{x} \quad (\text{付録 B.3})$$

■2変数に対する回帰方程式 (重回帰解析)

データ y を, 2 個のデータ (x_1, x_2) を用いて表現する回帰方程式

$$y(x_1, x_2) = a_1x_1 + a_2x_2 + b \quad (\text{付録 B.4})$$

の係数は次のように求められる.

$$a_1 = \frac{\begin{vmatrix} S_{x_1y} & S_{x_1x_2} \\ S_{x_2y} & S_{x_2x_2} \end{vmatrix}}{\begin{vmatrix} S_{x_1x_1} & S_{x_1x_2} \\ S_{x_1x_2} & S_{x_2x_2} \end{vmatrix}}, \quad a_2 = \frac{\begin{vmatrix} S_{x_1x_1} & S_{x_1y} \\ S_{x_1x_2} & S_{x_2y} \end{vmatrix}}{\begin{vmatrix} S_{x_1x_1} & S_{x_1x_2} \\ S_{x_1x_2} & S_{x_2x_2} \end{vmatrix}}, \quad b = \bar{y} - a_1\bar{x}_1 - a_2\bar{x}_2 \quad (\text{付録 B.5})$$

$$\text{ただし, } \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$