

ランダムネスの検定による傾向分析

帝京平成大学 鈴木義一郎

§ 1 逐次差の符号によるランダムネスの検定

ある講座での出席者数を1回目から10回目まで数えてみたら

講義回数	1	2	3	4	5	6	7	8	9	10
出席者数	128	113	110	100	108	101	104	98	92	82

のよ
うな結果が得られたとする。
これより、出席者数は減少傾向にあるといえるだろうか？

隣あった回数の中で、出席者数が減少していれば「0」、増加していれば「1」をつけてみると、(4,5)と(6,7)のところだけが1で、残りの7か所では0である。

一般に、

$$x_1, x_2, \dots, x_n$$

という観測系列が与えられたときに

$$Y(n) = \sum_{i=1}^{n-1} u(x_{i+1} - x_i) \quad , \quad u(t) = \begin{cases} 1 (t > 0) \\ 0 (t \leq 0) \end{cases}$$

といった統計量を定義すると、増加傾向にある個数を示す値である。

$Y(n)$ の期待値と分散は

$$E\{Y(n)\} = (n-1)/2 \quad , \quad V\{Y(n)\} = (n+1)/12$$

のように与えられる。さらに、 $D_n(k)$ を $\{1, 2, \dots, n\}$ の置換の中で $Y(n) = k$ となる組合せの数とすると、

$$D_{n+1}(0) = 1 \quad , \quad D_{n+1}(n) = 1, \\ D_{n+1}(k) = (n+1-k)D_n(k-1) + (k+1)D_n(k) \quad (1 \leq k \leq n-1)$$

といった漸化関係の成り立つことが分かる。 n 個の観測系列がランダムであるという仮説の下では、確率の値が

$$Pr\{Y(n) = k\} = D_n(k) / n! \quad (k = 0, 1, \dots, n-1)$$

のように与えられる。

したがって

$$\Pr\{Y(n) = k\} = D_n(k) / n! \quad (k = 0, 1, \dots, n-1)$$

のように与えられる。さて、
 $\Pr\{Y(n) \leq y(n, \alpha)\} \leq \alpha$

となる $y(n, \alpha)$ が表 1 のように与えられる。

例題の場合には、 $X(10) = 2$ であるから、有意水準 2.5% でランダムネスの仮説が棄却でき、出席者数は減少傾向にあると判断できる。

```

【Jによるプログラム】
stat_line=:[:+/:<}.
more=:[:+/"1[:*/[:>[:(,.,|.)&.>(1+[:i.>:@#);],0:
point_line=:4 :'+/"1 x.>:/+/%(more^(y.-1)1)!y.'
table_line=:3 :'.0.01 0.025 0.05 point_line"1 0 y.'

```

(}<}.)X=:128 113 110 100 108 101 104 98 92 82 0 0 0 1 0 1 0 0 0				stat_line X 2																							
]A=: (1+[:i.>:@#)1 1 1 2 3]B=: (],0:)1 1 1 1 0]C=: (,., .)A 1 3 2 2 3 1]D=: (,., .)B 1 0 1 1 0 1		C*D 1 0 2 2 0 1																					
]AB=: ((1+[:i.>:@#);],0:)1 1 <table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td>1</td><td>2</td><td>1</td><td>1</td></tr> <tr><td>3</td><td></td><td>0</td><td></td></tr> </table>		1	2	1	1	3		0]CD=: (,., .)&.>AB <table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td>1</td><td>3</td><td>1</td><td>0</td></tr> <tr><td>2</td><td>2</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>3</td><td>0</td><td>1</td></tr> </table>		1	3	1	0	2	2	1	1	1	3	0	1	*/> CD 1 0 2 2 0 1		+/"1*/> CD 1 4 1	
1	2	1	1																								
3		0																									
1	3	1	0																								
2	2	1	1																								
1	3	0	1																								
more 1 1 1 4 1 more^:2(1 1) 1 11 11 1		more^:3(1 1) 1 26 66 26 1 more^:4(1 1) 1 57 302 302 57 1		more^:5(1) 1 57 302 302 57 1 more^:6(1) 1 120 1191 2416 1191 120 1		+/more^:5(1) 720 +/more^:6(1) 5040																					
(i. 10), :more^:9(1)						+/more^:9(1)																					
0 1 2 3 4 5 6 7 8 9						3628800																					

1	1013	47840	455192	1310354	1310354	455192	47840	1013
1								

【%点の導出】

0.01	0.025	0.05	point_line	table_line 6+i.6	table_line 12+i.6
8				1 1 1	3 4 4
2 2 2				1 2 2	4 4 4
0.01	0.025	0.05	point_line	2 2 2	4 4 5
9				2 2 3	4 5 5
2 2 3				2 3 3	5 5 6
0.01	0.025	0.05	point_line	3 3 3	5 6 6
10					
2 3 3					

表1 ランダムネスの検定のパーセント点 $y(n, \alpha)$ の表

n	1%	2.5%	5%	n	1%	2.5%	5%	n	1%	2.5%	5%
5	1	1	1	18	6	6	6	31	11	12	12
6	1	1	1	19	6	7	7	32	12	12	13
7	1	2	2	20	7	7	7	33	12	13	13
8	2	2	2	21	7	7	8	34	13	13	14
9	2	2	3	22	7	8	8	35	13	14	14
10	2	3	3	23	8	8	9	36	13	14	15
11	3	3	3	24	8	9	9	37	14	15	15
12	3	4	4	25	9	9	10	38	14	15	16
13	4	4	4	26	9	10	10	39	15	15	16
14	4	4	5	27	10	10	11	40	15	16	16
15	4	5	5	28	10	11	11	45	18	18	19
16	5	5	6	29	10	11	11	50	20	21	21
17	5	6	6	30	11	11	12	60	24	25	26

§ 2 全ての対の符号によるランダムネスの検定

観測系列の $i < j$ というペア (X_i, X_j) に対して、 $X_i < X_j$ なら「1」、 $X_i \geq X_j$ なら「0」を

与えて、

$$Z(n) = \sum_{i=1}^{n-1} \sum_{j>i}^n u(x_j - x_i) \quad , \quad u(t) = \begin{cases} 1 (t > 0) \\ 0 (t \leq 0) \end{cases}$$

といった統計量が定義できて、これまた増加傾向にある個数を示す値で、0 から ${}_n C_2 = n(n-1)/2$ の範囲の値をとり、 $Z(n)$ の期待値と分散は

$$E\{Z(n)\} = n(n-1)/4 \quad , \quad V\{Z(n)\} = n(n-1)(2n+5)/72$$

のように与えられる。さらに、 $E_n(k)$ を $\{1, 2, \dots, n\}$ の置換の中で $Z(n) = k$ となる組合せの数とすると、

$$\Pr\{Z(n) = k\} = E_n(k)/n! \quad (k = 0, 1, \dots, n-1)$$

例えば $n = 4$ の場合、 $E_4(0)$ から $E_4(6)$ までの値が $\{1, 3, 5, 6, 5, 3, 1\}$ だから

$$\Pr\{Z(4) = 0\} = \Pr\{Z(4) = 6\} = 1/24 \quad , \quad \Pr\{Z(4) = 1\} = \Pr\{Z(4) = 5\} = 3/24 = 1/8$$

$$\Pr\{Z(4) = 2\} = \Pr\{Z(4) = 4\} = 5/24 \quad , \quad \Pr\{Z(4) = 3\} = 6/24 = 1/4$$

といった具合に算出される。

一般に

$$\Pr\{Z(n) \leq p(n, \alpha)\} \leq \alpha$$

となるような $p(n, \alpha)$ が表 2 のように与えられる。

表 2 ランダムネスの検定のパーセント点 $p(n, \alpha)$ の表

n	1%	2.5%	5%	n	1%	2.5%	5%	n	1%	2.5%	5%
5	0	0	1	18	45	50	54	31	164	174	183
6	1	1	2	19	52	57	61	32	176	187	196
7	2	3	4	20	59	64	69	33	188	200	210
8	4	5	6	21	66	72	78	34	202	214	224
9	6	8	9	22	74	80	85	35	215	228	239
10	9	11	12	23	82	89	94	36	229	242	254
11	12	14	16	24	91	98	104	37	244	257	269
12	15	18	20	25	100	107	114	38	259	273	285
13	19	22	25	26	109	117	124	39	274	289	301
14	24	27	29	27	119	128	135	40	290	305	318
15	28	32	35	28	130	139	146	45	376	394	410
16	34	37	41	29	140	150	158	50	473	495	513
17	39	43	47	30	152	162	170	60	702	731	755

さて与えられた例題の場合には

stat_pair X

4

であるから、有意水準 1%でもランダムネスの仮説が棄却できる。

```
【Jによるプログラム(直接計算)】
stat_pair=:[:+/([:+/"1{.<}.)¥.
distribution=:[:+/"1[:=:stat_pair"1([:i.! )A.i.
```

<pre>{.<}.)X 0 0 0 0 0 0 0 0 {.<}.)2}.X 0 0 0 0 0 0 0 0 {.<}.)3}.X 1 1 1 0 0 0 0 0 {.<}.)4}.X 0 0 0 0 0 0 0 0 {.<}.)5}.X 1 0 0 0 0 0 0 0</pre>	<pre>([:+/{.<}.)X 0 [:+/{.<}.)2}.X 0 [:+/{.<}.)3}.X 3 [:+/{.<}.)4}.X 0 [:+/{.<}.)5}.X 1</pre>	<pre>(({.<}.)¥.)X 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0</pre>	
<pre>([:+/:+/"1{.<}.)¥.)X 4</pre>		<pre>stat_pair X 4</pre>	
<pre>]a=:0 A. i.3 0 1 2</pre>	<pre>]b=:1 A. i.3 0 2 1</pre>	<pre>stat_pair a 3</pre>	<pre>stat_pair b 2</pre>
<pre>]c=:(([:i.!)A.i.)3 0 1 2 0 2 1 1 0 2 1 2 0 2 0 1 2 1 0</pre>	<pre>]d=:stat_pair"1 c 3 2 2 1 1 0 = d 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1</pre>	<pre>+/"1 = d 1 2 2 1]D3=:distribution 3 1 2 2 1]D4=:distribution 4 1 3 5 6 5 3 1</pre>	

]D5=:distribution 5 1 4 5 9 15 20 22 20 15 9 4 1]D6=:distribution 6 1 5 14 29 49 71 90 101 101 90 71 49 29 14 5 1
6!:2' D7=:distribution 7' 0.413477 6!:2' D9=:distribution 9' 33.7535 , \$&> D6;D7;D8;D9;D10 16 22 29 37 46	6!:2' D8=:distribution 8' 3.7184 6!:2' D10=:distribution 10' 696.271 【直接計算するのは時間がかかりすぎる！】

【Jによるプログラム(試行錯誤による漸化関係の利用)】

```

next=:+:@#-[+:[i.[:<:no=[:-[:>[:%:(8:*#)-7:
half=#-[:<[::-no1=[+:[i.[:<:no
take=[+:/half{.}
tail=:13 :'-/((-,) (#u)->:no y.) {.'0 1 u=.take y.'
index=:_1:-[:i.>:@no-half
modify=:13 :'(tail y.)(index y.))take y.'
mdfy=:+/¥`modify@.(#>2:)
more_pair=:3 : 'u,|.((next-half)y.) {u=.mdfy y.'
point_pair=:4 : '<+/'1 x.>:/(+/¥more_pair^(y.-1)1)%!y.'

```

<pre>no L:0 D3;D4;D5;D6;D7;D8;D9;D10 3 4 5 6 7 8 9 10 next L:0 D3;D4;D5;D6;D7;D8;D9;D10 7 11 16 22 29 37 46 56 \$ L:0 D4;D5;D6;D7;D8;D9;D10 7 11 16 22 29 37 46</pre>	<p>求めようとしている分布の“標本数”を与える補助関数</p> <p>「next」は次に求めようとする分布の“範囲”を与えている。</p> <p>「distribution」によって求めたD4~D10の分布の“範囲”を与えている。</p>
<pre>\$_@take L:0 D3;D4;D5;D6;D7;D8;D9;D10 4 6 8 11 15 19 23 28 \$_@modify L:0 D3;D4;D5;D6;D7;D8;D9;D10 4 6 8 11 15 19 23 28 >.@- :@\$ L:0 D4;D5;D6;D7;D8;D9;D10 4 6 8 11 15 19 23 no1 L:0 D3;D4;D5;D6;D7;D8;D9;D10 1 3 6 10 15 21 28 36</pre>	<p>「\$_@take」は次に求めようとする分布の範囲の“半分”の数を与えている。</p> <p>「modify」の出力結果と「take」の出力結果の長さは同じである。</p> <p>「distribution」によって求めたD4~D10の分布の“範囲”の半分を与えている。</p> <p>「take」で与えた数値の末尾の“何個”を修正するか、その個数を与えている。</p>

<pre>]t4=(3 :'((-,])(<:@half-no)y.){."0 1 u=.take y.')}D4 23 1]t5=(3 :'((-,])(<:@half-no)y.){."0 1 u=.take y.')}D5 91 106 1 5</pre>	<p>-/ t4 22 「tail」の値が出力される -/ t5 90 101 D6の「tail」の値を出力</p>								
<pre>(index;tail;take;modify)D4</pre> <table border="1" data-bbox="252 1541 778 1617"> <tr> <td>_1</td> <td>22</td> <td>1 4 9 15 20</td> <td>1 4 9 15 20</td> </tr> <tr> <td></td> <td></td> <td>23</td> <td>22</td> </tr> </table>	_1	22	1 4 9 15 20	1 4 9 15 20			23	22	<p>「modify」は「take」の出力結果の「index」の位置の要素を「tail」の数で置き換える関数である。（次の分布の半分が完成）</p>
_1	22	1 4 9 15 20	1 4 9 15 20						
		23	22						
<pre>(index;tail;take;modify)D5</pre> <table border="1" data-bbox="252 1684 1129 1751"> <tr> <td>_2 _1</td> <td>90 101</td> <td>1 5 14 29 49 71 91</td> <td>1 5 14 29 49 71 90 101</td> </tr> <tr> <td></td> <td></td> <td>106</td> <td></td> </tr> </table>	_2 _1	90 101	1 5 14 29 49 71 91	1 5 14 29 49 71 90 101			106		<p>D6の分布の半分が完成する。</p>
_2 _1	90 101	1 5 14 29 49 71 91	1 5 14 29 49 71 90 101						
		106							

<pre> .. D4:more_pair D3 1 3 5 6 5 3 1 1 3 5 6 5 3 1 D4 -: more_pair D3 1 </pre>	<pre> .. D5:more_pair D4 1 4 9 15 20 22 20 15 9 4 1 1 4 9 15 20 22 20 15 9 4 1 D5 -: more_pair D4 1 </pre>	<p>「more_pair」はDkからDk+1を出力する関数である。</p> <p>「more_pair D3」とD4そして「more_pair D4」とD5は完全に一致している。</p>
<pre> .. D6:more_pair D5 1 5 14 29 49 71 90 101 101 90 71 49 29 14 5 1 1 5 14 29 49 71 90 101 101 90 71 49 29 14 5 1 D6 -: more_pair D5 1 </pre>		<p>「more_pair D5」とD6を同時に表示している。</p> <p>「more_pair D5」とD6は完全に一致している。</p>
<pre> .. D7:more_pair D6 1 6 20 49 98 169 259 359 455 531 573 573 531 455 359 259 169 98 49 20 6 1 1 6 20 49 98 169 259 359 455 531 573 573 531 455 359 259 169 98 49 20 6 1 D7 -: more_pair D6 1 </pre>		
<pre> D8 -: more_pair D7 1 </pre>	<pre> D9 -: more_pair D8 1 </pre>	<pre> D10 -: more_pair D9 1 </pre>

【%点の導出】

<pre> (p=:0.01 0.025 0.05) point_pair 5 0 0 1 p point_pair 6 1 1 2 p point_pair 7 2 3 4 p point_pair 8 4 5 6 </pre>	<pre> p point_pair"1 0(6+i.10) 1 1 2 2 3 4 4 5 6 6 8 9 9 11 12 12 14 16 15 18 20 </pre>
---	---

p point_pair 9 6 8 9	19 22 25 24 27 29 28 32 35
-------------------------	----------------------------------

【Jによるプログラム(近似計算)】

```

var_pair=(5:+2:*)]*<:)%72"_
forth=:3 :'(y.*(y.-1)*_372_997_127_328_100&p.y.)%43200'"0
range_pair=([:(:)--@(:)][:i.1:+[::-]*<:)]%[:var_pair
nden=([:^[:-*:%2:)]%[:[:o.2:
middle=[:].]*0.002*i.500)"_
ndf=:0.5"+ ndfs=:]*([+/([::-[:nden 0:]),[:nden middle)%500"_
orgf=:nden*0_3 0 1&p.
approximation=:ndf@[-orgf@[*(_3:+forth%[:*var_pair)]@]%24"_
modify=:13 :'(tail y.)(_1-i.3-no y.))take y.'
point_ap=[:<[:+/"1]>:/[:(range_pair approximation"0)]

```

0.01 0.025 0.05 point_ap 10 9 11 13(9 11 12)	0.01 0.025 0.05 point_ap 40 290 306 319(290 305 318)
0.01 0.025 0.05 point_ap 20 59 64 69	0.01 0.025 0.05 point_ap 50 474 495 514(473 495 513)
0.01 0.025 0.05 point_ap 30 152 162 171(152 162 170)	0.01 0.025 0.05 point_ap 60 703 731 755(702 731 755)

【東京地区平成7年の最低気温のデータ】

```

MIN2=:2.1 1.5 2.9 3.0 3.3 3.7 4.7 4.2 4.8 4.9
MIN2=:MIN2,3.0 2.9 1.1 1.8 2.0 3.7 5.0 3.3 0.5 0.9
MIN2=:MIN2,0.5 _1.4 _0.4 2.3 3.9 8.4 7.2 6.7
MIN3=:6.7 4.6 5.2 4.0 5.0 6.7 13.4 9.1 6.6 4.6
MIN3=:MIN3,7.7 6.1 4.6 9.2 6.2 4.3 3.5 5.6 6.9 8.0
MIN3=:MIN3,6.3 8.4 5.4 4.2 3.0 10.4 8.4 6.3 12.0 10.4 7.9
MIN4=: 8.0 11.1 10.5 10.7 13.3 12.6 10.9 10.3 12.9 12.6
MIN4=:MIN4,10.5 8.8 9.5 12.0 10.6 8.9 9.9 11.2 11.0 12.1
MIN4=:MIN4,13.5 14.5 9.9 9.6 11.1 11.8 11.3 14.1 13.2 16.1

```

stat_line MIN2	stat_line MIN3	stat_line MIN4
16	12	15

stat_pair MIN2 194	stat_pair MIN3 268	stat_pair MIN4 269
-----------------------	-----------------------	-----------------------