

『連勝の確率』について

統計数理研究所(名誉教授) 鈴木義一郎

§1 成功連の長さの確率

戦績の系列で、1の表れる確率をp(0の表れる確率をq=1-p)とすると、全ての系列のなかで、1という文字がrまでしか連続しないような系列の集合をS(r,n)と表す。さらにS(r,n)のなかで、1が丁度h回、0が(n-h)回表れる系列の集合をT(h,r,n)、その個数をC(h,r,n)とする。

「S(r,n)という系列を得る確率」Q(p;r,n)は

Q(p;r,n) = sum_{h=0}^n C(h;r,n) p^h q^{n-h}

のように与えられる。したがって、1-Q(p;r,n)が(r+1)以上の連勝の表れる確率を与えることになる。

さて係数の間には、次のような“漸化関係”が成り立つ。

C(h;r,n) = sum_{i=0}^r C(h-1;r,n-i-1)

C(0;r,n) = 1, C(1;r,n) = r, C(h;r,n) = n C_h (0 <= h < r)

Table with 2 columns: code snippets and explanatory text. The code defines variables for calculating C(h;r,n) and Q(p;r,n). The text explains the use of the recurrence relation and the definition of the probability function.

]C=:first 2 6]D=:second C	change D	change^:2 D																										
1 0 0	<table border="1"><tr><td>0 0 0</td><td>3 3 0</td></tr><tr><td>1 0 0</td><td>4 6 0</td></tr><tr><td>2 1 0</td><td>5 10 0</td></tr><tr><td></td><td>6 15 0</td></tr></table>	0 0 0	3 3 0	1 0 0	4 6 0	2 1 0	5 10 0		6 15 0	<table border="1"><tr><td>0 0 0</td><td>4 6 0</td></tr><tr><td>1 0 0</td><td>5 10 0</td></tr><tr><td>2 1 0</td><td>6 15 0</td></tr><tr><td>3 3 0</td><td></td></tr></table>	0 0 0	4 6 0	1 0 0	5 10 0	2 1 0	6 15 0	3 3 0		<table border="1"><tr><td>0 0 0</td><td>5 10 0</td></tr><tr><td>1 0 0</td><td>6 15 0</td></tr><tr><td>2 1 0</td><td></td></tr><tr><td>3 3 0</td><td></td></tr><tr><td>4 6 2</td><td></td></tr></table>	0 0 0	5 10 0	1 0 0	6 15 0	2 1 0		3 3 0		4 6 2	
0 0 0	3 3 0																												
1 0 0	4 6 0																												
2 1 0	5 10 0																												
	6 15 0																												
0 0 0	4 6 0																												
1 0 0	5 10 0																												
2 1 0	6 15 0																												
3 3 0																													
0 0 0	5 10 0																												
1 0 0	6 15 0																												
2 1 0																													
3 3 0																													
4 6 2																													
1 2 1																													
1 3 3																													
1 4 6																													
1 5 10																													
1 6 15																													

slide C	0 line 2 6	4 line 2 6	(i.7)line"0 1(2 6)
0 0 0	1 1 1 1 1 1	0 0 0 0 1 6	1 0 0 0 0 0
1 0 0	1 line 2 6	5 line 2 6	1 1 0 0 0 0
2 1 0	0 1 2 3 4 5 6	0 0 0 0 0 0	1 2 1 0 0 0
3 3 0	2 line 2 6	6 line 2 6	1 3 3 0 0 0
4 6 2	0 0 1 3 6 10 15	0 0 0 0 0 0	1 4 6 2 0 0
5 10 7	3 line 2 6		1 5 10 7 1 0 0
6 15 16	0 0 0 0 2 7 16		1 6 15 16 6 0 0
:coef 2 6	coef 3 6	coef 4 6	
1 0 0 0 0 0	1 0 0 0 0 0	1 0 0 0 0 0	
1 1 0 0 0 0	1 1 0 0 0 0	1 1 0 0 0 0	
1 2 1 0 0 0	1 2 1 0 0 0	1 2 1 0 0 0	
1 3 3 0 0 0	1 3 3 1 0 0	1 3 3 1 0 0	
1 4 6 2 0 0	1 4 6 4 0 0	1 4 6 4 1 0 0	
1 5 10 7 1 0 0	1 5 10 10 3 0 0	1 5 10 10 5 0 0	
1 6 15 16 6 0 0	1 6 15 20 12 2 0	1 6 15 20 15 4 0	

]C3=:coef 2 3]P3=:0.5 prob"0 i.>:3]CP3=:C3*P3	+/"1 CP3
	1 0 0 0	1 0 0 0	1 1 1 0.875
1 0 0 0	0.5 0.5 0 0	0.5 0.5 0 0	0.5 probr 2 3
1 1 0 0	0.25 0.25 0.25 0	0.25 0.5 0.25 0	1 1 1 0.875
1 2 1 0	0.125 0.125 0.125 0.125	0.125 0.375 0.375 0	
1 3 3 0			

p	$S(r, n)$ という系列を得る確率 $Q(p; r, n)$ を出力する関数		
0.5	0.5	0.5	0.5
0.448125	0.834667	0.95854	0.990016
0.6	0.6	0.6	0.6
0.164964	0.574411	0.833851	0.940716

下表に、成功連（1の連）の長さが3以内の場合で、成功の確率を $p=0.5, 0.6, 0.7, 0.8, 0.9$ とした場合の確率を示してみた。これらの値を1から引いた確率 p の1人が4連勝以上する確率を与える。 p の値が大きくなるほど、また勝負回数 n が多くなるほど連勝の確率が大きくなる様子が読みとれる。

確率の値が異なる場合の3連勝以内の確率(%表示)

n	p=0.5	p=0.6	p=0.7	p=0.8	p=0.9
10	75.4883	57.5484	36.7497	17.2700	4.0638
11	72.7539	53.5738	32.3138	13.7756	2.6684
12	70.1172	49.8678	28.3968	10.9523	1.7035
13	67.5781	46.4306	24.9985	8.8001	1.1690
14	65.1306	43.2273	21.9946	7.0441	0.7826
15	62.7716	40.2440	19.3475	5.6294	0.5160
16	60.4980	37.4668	17.0199	4.5009	0.3409
17	58.3069	34.8816	14.9745	3.6036	0.2291
18	56.1951	32.4746	13.1738	2.8827	0.1524
19	54.1597	30.2337	11.5896	2.3057	0.1011
20	52.1981	28.1475	10.1960	1.8445	0.0672

$r, 6! : 2^r = 0.5$ prob $1/3$ 20'	$r, 6! : 2^r = 0.5$ prob $1/3$ 150'	勝負回数 n が増えても原理的には計算が可能であるが、計算時間が膨大になる。
0.6875 0.226292	0.027285 93.6612	
$r, 6! : 2^r = 0.5$ prob $1/3$ 50'	$r, 6! : 2^r = 0.5$ prob $1/3$ 200'	そこで、近似式を考えてやる方が合理的になる。
0.249587 2.96026	0.00431361 240.842	
$r, 6! : 2^r = 0.5$ prob $1/3$ 100'	$r, 6! : 2^r = 0.5$ prob $1/3$ 300'	
0.0394584 26.2172	0.172586 1038.16 (17分強)	

§2 成功連の長さの確率に対する近似式

§1 で示した方法は、 n が大きくなっても原理的には計算可能であるが、 n が 200 程度になると計算時間が 17 分以上もかかるので、実際的とはいえなくなる。そこで、近似式による求める方法を考えてみよう。

n 回の勝負で丁度 r の成功連が生じているという事象の確率を u_r とすると、

$$u_n = 0 \quad (0 \leq n < r) \quad , \quad u_r = p^r \quad , \quad u_{r+1} = (1-p)p^r$$

であるから

$$u_{r+1} + u_r p = p^r$$

という関係が成り立つ。

さらに一般に、 $n \geq r$ に対しても

$$u_n + u_{n-1}p + u_{n-2}p^2 + \cdots + u_{n-r+1}p^{r-1} = p^r$$

という関係の成り立つことも分かる。両辺に s^n を掛けて $n = r, r+1, \cdots$ について加えると、左辺は

$$\begin{aligned} & \sum_{n=1}^{\infty} u_n s^n + ps \sum_{n=r}^{\infty} u_{n-1} s^{n-1} + \cdots + p^{r-1} s^{r-1} \sum_{n=r}^{\infty} u_{n-r+1} s^{n-r+1} \\ &= \sum_{n=1}^{\infty} u_n s^n + ps \sum_{n=1}^{\infty} u_n s^n + \cdots + p^{r-1} s^{r-1} \sum_{n=1}^{\infty} u_n s^n \\ &= \{U(s) - 1\} \{1 + ps + p^2 s^2 + \cdots + p^{r-1} s^{r-1}\} \\ &= \{U(s) - 1\} \frac{1 - p^r s^r}{1 - ps} \end{aligned}$$

のように表される。ここで

$$U(s) = \sum_{n=0}^{\infty} u_n s^n \quad (u_0 = 1)$$

である。また右辺も

$$p^r \sum_{n=r}^{\infty} s^n = \frac{p^r s^r}{1 - s}$$

のように表されるから

$$(+)\quad U(s) = \frac{1 - s + (1 - p)p^r s^{r+1}}{(1 - s)(1 - p^r s^r)}$$

次に、長さ r の成功連が、 n 回目で初めて生じているという事象の確率を v_r とすると、

明らかに

$$(*)\quad v_n = 0 \quad (0 \leq n < r) \quad , \quad v_r = p^r \quad , \quad v_{r+1} = (1 - p)p^r$$

である。さらに

$$u_n = \sum_{h=1}^n v_h u_{n-h}$$

という関係も成り立つ。そこで

$$V(s) = \sum_{n=0}^{\infty} v_n s^n$$

という関数を定義すると、(*)の関係より

$$\sum_{n=1}^M u_n s^n = \sum_{n=1}^M s^n \sum_{h=1}^n v_h u_{n-h} = \sum_{h=1}^M v_h s^h \sum_{n=h}^M u_{n-h} s^{n-h} = \sum_{h=1}^M v_h s^h \sum_{m=0}^{M-h} u_m s^m$$

$M \rightarrow \infty$ とすると

$$U(s) - 1 = V(s)U(s)$$

ところで(+)の関係があったから

$$\begin{aligned} V(s) &= 1 - \frac{1}{U(s)} = \frac{(1-s + (1-p)p^r s^{r+1}) - (1-s)(1-p^r s^r)}{1-s + (1-p)p^r s^{r+1}} = \frac{p^r s^r (1-ps)}{1-s + (1-p)p^r s^{r+1}} \\ &= \frac{p^r s^r}{(1-s + (1-p)p^r s^{r+1})/(1-ps)} = \frac{p^r s^r}{(1-p)s\{1 + 2ps + \dots + p^{r-1}s^{r-1}\}} \equiv \frac{A(s)}{B(s)} \end{aligned}$$

のように表される。さらに分母の関数 $B(s)$ は s に関する r 次の多項式で、しかも

$$C(s) = (1-p)\{1 + ps + \dots + rp^{r-1}s^{r-1}\}$$

という関数が $s \geq 0$ で 0 から ∞ まで単調に増加し

$$C(1) = (1-p)\{1 + p + \dots + rp^{r-1}\} = (1-p^r) < 1$$

であるから、 $B(s)$ は唯一個の正の実根 $s_0 (> 1)$ を持つことが分かる。さらに微分も

$$B^{(1)}(s) = -(1-p)\{1 + 2ps + \dots + rp^{r-1}s^{r-1}\}$$

のように与えられるから

$$\rho = \frac{-A(s_0)}{B^{(1)}(s_0)} = \frac{p^r s_0^r}{(1-p)\{1 + 2ps_0 + \dots + rp^{r-1}s_0^{r-1}\}}$$

のように算出できる。代数学の基本定理により $U(s) = A(s)/B(s)$ の s^n の係数 u_n に対して

$$u_n = \frac{\rho}{s_0^{n+1}} = \frac{p^r s_0^r}{(1-p)\{1 + 2ps_0 + \dots + rp^{r-1}s_0^{r-1}\}} \cdot \frac{1}{s_0^{n+1}}$$

といった近似関係の成り立つことが示される。従って

$$f_n = u_{n+1} + u_{n+2} + u_{n+3} + \dots = \frac{p^r s_0^r / (1-s_0)}{(1-p)\{1 + 2ps_0 + \dots + rp^{r-1}s_0^{r-1}\}} \cdot \frac{1}{s_0^{n+1}}$$

この f_n は n までに長さ r の成功連が生じていないという事象の確率を表わすから、先の議

論の r を $(r+1)$ に変えて、 1 という文字が r までしか継続しないような集合 $S(r, n)$ の系

列を得る確率 $Q(p; r, n)$ に対する近似式は

$$q_n = \frac{p^{r+1} x^{r+1} / (1-x)}{(1-p) \{1 + 2px + \dots + rp^r x^r\}} \cdot \frac{1}{x^{n+1}}$$

で与えられる。ここで x は、 $(r+1)$ 次の方程式

$$1 - (1-p)x \{1 + px + \dots + rp^r x^r\} = 0$$

の唯一の実根である。

<pre> prob_ap=:4 :0 q=.1,(x-1)*x^i.r=>{:y s=(1-q&p.%q&p.D.1)^:50(1) (1-x*s)%((1-x)*1+r- r*s)*s^>{:y) </pre>	<pre>]q=:0.6(pol=:4 '1,(x-1)*x^i.>:y')3 1_0.4_0.24_0.144_0.0864 root=:4:'(1-q&p.(q=.x pol y)&p.D.1)^:50(1)']s=:0.8 root 3 q&p. s 1.25 1.11022e_16 </pre>	
	<p>「prob_ap」が $S(r, n)$ という系列を得る確率 $Q(p; r, n)$ を出力する関数である。</p>	

近似式による確率の値は正確な確率とほとんど遜色なし！			
0.5 prob1 4 50 0.448125	0.5 prob1 6 50 0.834667	0.5 prob1 8 50 0.95854	0.5 prob1 10 50 0.990016
0.6 prob1 4 50 0.164964	0.6 prob1 6 50 0.574411	0.6 prob1 8 50 0.833851	0.6 prob1 10 50 0.940716

0.5 prob_ap 4 50 0.448125	0.5 prob_ap 6 50 0.834667	0.5 prob_ap 8 50 0.95854	0.5 prob_ap 10 50 0.990016
0.6 prob_ap 4 50 0.164964	0.6 prob_ap 6 50 0.574411	0.6 prob_ap 8 50 0.833851	0.6 prob_ap 10 50 0.940716

```

1-(0.9+0.01*i.6)prob_ap"0 1(69 900)
0.0514843 0.0984685 0.181174 0.315425 0.506851 0.727632

```

```
a=. (0.9+0.01*i.6)
```

```
a., 1- (0.9+0.01*i.6) prob_ap ("0 1) 69 900
0.9 0.0514843
0.91 0.0984685
0.92 0.181174
0.93 0.315425
0.94 0.506851
0.95 0.727632
```

900(=90×10)戦の中で70連勝以上できるには、勝率 p が0.93以上が必要である！