N IFUSP

Instituto de Física
Universidade de São Paulo

O COMPRIMENTO MÁXIMO ESPERADO PARA AS CADEIAS DE CARAS CONSECUTIVAS EM “N"LANÇAMENTOS DE UMA MOEDA, SENDO "P"A PROBABILIDADE DE CARA E "Q" A DE COROA É DADO POR "- LOGpNQ"

BORGE, CJ

Publicação IF – 1604/2005

UNIVERSIDADE DE SÃO PAULO
Instituto de Física
Cidade Universitária
Caixa Postal 66.318
05315-970 - São Paulo - Brasil

O comprimento máximo esperado para as cadeias de caras consecutivas em "n" lançamentos de uma moeda, sendo "p" a probabilidade de cara e "q" a de coroa é dado por " $-\log_p nq$ "

Carlos J. Borge

Introdução

Em 1999 Ted Hill [1] afirmou que um dado professor pede a seus alunos que lancem uma moeda por 200 vezes e anotem os resultados. Ao olhar para a seqüência de caras e coroas que determinado aluno lhe entrega, ele sabe se o aluno lançou realmente a moeda ou se inventou os resultados. O professor apenas verifica qual é o comprimento da maior cadeia de caras consecutivas e se este comprimento não estiver em torno de seis, então ele conclui que o aluno, muito provavelmente inventou os resultados.

Como o professor pode saber que a maior cadeia de caras terá comprimento próximo de seis?

É exatamente esta pergunta que desejamos responder. E a resposta é bem simples!

Usando o logaritmo do título calculamos:

$$-\log_{1/2}(200 \times ½) = 6{,}64$$

Indicando que a maior cadeia de caras consecutivas terá comprimento próximo desse valor.

E a probabilidade de o referido comprimento estar em torno desse logaritmo é muito alta, como veremos adiante.

Mais genericamente, se lançarmos a moeda "n" vezes e a probabilidade de cara for "p", invariável ao longo dos lançamentos, e a de coroa for "q", também invariável, sendo:

$$p + q = 1, \quad 0 < p < 1, \quad (0 < q < 1) \text{ e } \quad nq > 1/p \quad (nq \text{ é o produto entre n e q}),$$

então, o referido comprimento máximo, que a partir de agora chamaremos de "$L_{p,MAX}$", é uma variável aleatória e deverá assumir um valor inteiro próximo de:

$$-\log_p nq.$$

Repare que os lançamentos da moeda definidos do modo acima, são exemplos dos conhecidos ensaios de Bernoulli, e que a probabilidade de cara não precisa ser igual a meio. Este logaritmo (respeitando-se as condições de validade) se aplica para qualquer exemplo de ensaios de Bernoulli, e não apenas para os lançamentos de uma moeda.

Descobri esse logaritmo de um modo intuitivo e não rigoroso, como veremos a seguir, mas desenvolvi um modo preciso de se calcular a sua probabilidade de ocorrência para

$$p = q = ½,$$

que o transforma numa poderosa ferramenta matemática.

Como surgiu o logaritmo

Vamos supor que você vai jogar uma moeda (honesta ou não) por $n = 1024$ vezes, e vamos chamar de "$L_p(k)$" o comprimento da cadeia de caras no instante "k", isto é, imediatamente após o k-ésimo lançamento da moeda. Então, "$L_p(k)$" vai assumir 1024 valores nesse exemplo, e "k" vai variar de 1 até 1024.

Vamos supor ainda que se o resultado for coroa, então neste instante $L_p(k)$ = zero. Se sair cara imediatamente após uma coroa, então neste instante $L_p(k) = 1$ (indicando o começo de uma cadeia de caras). E se sair outra cara, logo após esta citada, então "$L_p(k)$" vai para 2 (indicando que neste momento a cadeia de caras está com comprimento igual a 2). Depois do valor 2, "$L_p(k)$" assumirá o valor 3 (se sair cara novamente) ou o valor zero (se sair coroa).

Consideremos o exemplo da tabela 1 em que uma moeda foi lançada por $n = 18$ vezes. Seja "S" a representação para resultados do tipo cara (sucesso nos ensaios de Bernoulli) com probabilidade "p", fixa, de ocorrência e "F" a representação para resultados do tipo coroa (fracasso nos ensaios de Bernoulli) com probabilidade "q", fixa, de ocorrência. E vamos acompanhar os valores de "k" e de "$L_p(k)$" para a seqüência de 18 resultados:

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Resultado	S	S	F	S	F	F	F	S	F	S	S	S	F	S	F	S	F	F
$L_p(k)$	1	2	0	1	0	0	0	1	0	1	2	3	0	1	0	1	0	0

Tabela1. Uma moeda foi lançada por 18 vezes e os resultados do tipo cara (S) e coroa (F) obtidos foram anotados, bem como o valor da nossa outra variável aleatória "$L_p(k)$" que indica o comprimento da cadeia de caras consecutivas após o k-ésimo lançamento. Exemplificando: no sétimo lançamento (k = 7) obtivemos coroa (F) e $L_p(7) = 0$. A maior cadeia de caras consecutivas ocorreu no décimo segundo lançamento (k = 12), e tem comprimento
$L_p(12) = L_{p,MAX} = 3$.

Com isso fica fácil entender de onde surgiu o logaritmo. Reforce a sua atenção agora.

Em 1024 lançamentos, se a probabilidade de cara for $p = ½$, então a de coroa também será igual a ½, e devemos esperar que aconteçam em torno de 512 coroas, o que implica em $L_p(k)$ igual a zero por aproximadamente 512 vezes.

Cada uma das coroas obtidas é seguida de cara ou de coroa, e devemos esperar que aproximadamente a metade delas seja seguida por cara, e assim "$L_p(k)$", nesses 1024 lançamentos, deverá assumir o valor 1 aproximadamente 256 vezes.

$L_p(k) = 2$, deve ocorrer aproximadamente 128 vezes.

$L_p(k) = 3$, deve ocorrer aproximadamente 64 vezes.

$L_p(k) = 4$, deve ocorrer aproximadamente 32 vezes.

$L_p(k) = 5$, deve ocorrer aproximadamente 16 vezes.

$L_p(k) = 6$, deve ocorrer aproximadamente 8 vezes.

$L_p(k) = 7$, deve ocorrer aproximadamente 4 vezes.

$L_p(k) = 8$, deve ocorrer aproximadamente 2 vezes.

$L_p(k) = 9$, deve ocorrer aproximadamente 1 vez.

E, então esperamos que a maior cadeia de resultados do tipo cara tenha comprimento próximo de 9.

Para esse exemplo o logaritmo vale:

$$- \log_p nq$$

$$- \log_{0,5} (1024 \times 0{,}5) = 9.$$

O logaritmo aparece da seguinte maneira:

Em n lançamentos da moeda (com n grande o suficiente) devemos esperar nq resultados do tipo coroa (F), e então $L_p(k)$ deve assumir o valor zero, por aproximadamente nq vezes.

Cada uma dessas nq coroas é seguida por cara ou coroa. E devemos esperar que aproximadamente nqp sejam seguidas por cara e nqq por coroa. Então, por nqp vezes $L_p(k)$ deve assumir o valor 1. (cara logo após coroa, indicando o início de uma cadeia de caras).

Se no instante r (i. é, após o r-ésimo lançamento), tivermos $L_p(r) = 1$, então $L_p(r + 1)$ valerá zero ou 2, e devemos esperar $L_p(r+1) = 2$ por aproximadamente nqp^2 vezes. E assim devemos esperar ainda:

$L_p(k) = 3$, deve ocorrer aproximadamente nqp^3 vezes.

$L_p(k) = 4$, deve ocorrer aproximadamente nqp^4 vezes.

$L_p(k) = 5$, deve ocorrer aproximadamente nqp^5 vezes.

.

.

.

$L_p(k) = x$, deve ocorrer aproximadamente nqp^x = uma vez.

Repare que x é o comprimento esperado da maior cadeia de caras consecutivas, e que devemos escolher $nqp^x = 1$ para que exista ao menos uma cadeia com o comprimento x.

Então: $nqp^X = 1$

$p^X = 1/nq$

$\log p^X = \log (1/nq)$

$x \log p = -\log nq$

$x = (-\log nq) / (\log p)$

$x = -\log_p nq.$

Probabilidade de ocorrência do logaritmo para p = q = ½

As tabelas 2 e 3 nos informam, para um número n de lançamentos da moeda, o valor de

$$x = -\log_p nq.$$

E também o número de possibilidades em que a maior cadeia de caras tem comprimento L igual a 0, 1, 2. 3, etc.

TABELA 2

TABELA 3

Por exemplo se pegarmos a linha correspondente a oito lançamentos da moeda (n = 8), perceberemos que a previsão é:

$$x = -\log_p nq.$$

$$x = 2,$$

e que a probabilidade exata de a maior cadeia de caras ter comprimento igual a 2, que chamaremos de P(2), é dada por:

$(94 / 256) = 0,3672$

isto é: $P(2) = 36,72\%$.

Portanto, tudo o que precisamos é construir a tabela.

O processo de construção dessa tabela é muito simples. Para zero lançamentos da moeda ($n = 0$), a maior cadeia de caras tem comprimento nulo e então aparece o número 1 na casa correspondente. (Significando: uma possibilidade para comprimento L igual a zero.)

Acima desse número 1, referido no parágrafo anterior, temos outro número 1 necessário para a lei de formação da tabela. Repare que as tabelas 2 e 3 são, na verdade, a mesma tabela. Essa tabela tem infinitas linhas e colunas, e na tabela 3 representamos uma linha a mais, apenas para melhorar a compreensão da lei de formação da mesma.

Vamos nomear as casas da tabela como cnL. Assim, c31 corresponde à casa em que $n = 3$ e $L = 1$, e seu conteúdo é o valor 4. A casa c(-1)0 contém o valor 1.

Para um único lançamento da moeda ($n = 1$), a maior cadeia de caras terá comprimento igual a 1 se sair cara e zero se sair coroa, então aparecem os números **1** e **1** nas casas c10 e c11 respectivamente. (**Uma** cadeia de comprimento zero e **uma** de comprimento um)

Se $n = 2$ (dois lançamentos da moeda) teremos na coluna zero (que corresponde a zero caras na maior cadeia de caras) o valor 1, que indica um único modo de se obter zero caras (F,F). Na coluna 1 (casa c21) teremos o valor 2 que corresponde a duas possibilidades de se ter em dois lançamentos da moeda, a maior cadeia de caras com comprimento igual a 1. (S,F ou F,S). E finalmente na coluna 2 (casa c22) teremos o número 1 que corresponde a uma possibilidade de termos duas caras em dois lançamentos da moeda (S,S).

A tabela 3 nos mostra, por exemplo, na linha 9 ($n = 9$) e coluna 2 (comprimento da maior cadeia de caras igual a 2), o valor 185, que significa o seguinte: "em nove lançamentos de uma moeda, se considerarmos todas as possibilidades para os nove resultados, teremos 185 seqüências em que a maior cadeia de caras tem comprimento igual a 2".

Para 9 lançamentos de uma moeda honesta, teremos 2^9 possibilidades para a seqüência de resultados, porém

$$2^9 = 512 = 1 + 88 + 185 + 127 + 63 + 28 + 12 + 5 + 2 + 1 \text{ (ver tabela 3).}$$

Cada um desses valores (1, 88, 185, ...) corresponde a um número de cadeias. Por exemplo: exatamente 185 dessas 512 cadeias, possuem a maior cadeia de caras com comprimento igual a 2. Ou então, em exatamente 127 das 512 seqüências possíveis, a maior cadeia de caras terá comprimento igual a 3.

Então, a probabilidade P(3), de a maior cadeia de caras ter comprimento igual a 3, em nove lançamentos de uma moeda honesta é exatamente igual a 127 / 512.

Repare que se quisermos comprimento máximo em torno de 2, isto é, 1, 2, ou 3, teremos uma probabilidade altíssima de ocorrência. $P(1) + P(2) + P(3) = (88 + 185 + 127) / 512$. (78 %)

Lei de formação da tabela

Se continuarmos nessa linha, correspondente a nove lançamentos da moeda ($n = 9$), na tabela 3, seguindo as cores, poderemos perceber que:

$88 = 54 + 33 + 1$;
$185 = 94 + 47 + 23 + 20 + 1$;
$127 = 59 + 27 + 12 + 5 + 11 + 12 + 1$;
etc.

O que nos dá a lei de formação da tabela.

Resultado de simulação

Simulamos $n = 65000$ lançamentos de uma moeda com probabilidade de cara $p = 1/5$ e de coroa $q = 4/5$.

A previsão, através do logaritmo, para o comprimento da maior cadeia de caras consecutivas é um número inteiro em torno de 6,75. Tivemos, como resultado da simulação, comprimento igual a 6.

Para as cadeias de coroas consecutivas, o logaritmo ($-\log_q np$) assume o valor 42,45 e o comprimento máximo medido foi 38.

Como podemos perceber a previsão foi boa.

Casos especiais

Para $p = q = ½$ o logaritmo se aplica muito bem, mas quando p se afasta muito de ½ e o número de ensaios é pequeno, a previsão pode não se adequar com a realidade.

Se, por exemplo, em onze ensaios tivermos $p = 9/10$ e consequentemente $q = 1/10$, um resultado com dez caras e uma coroa terá boa chance ocorrer.

Uma coroa distribuída ao acaso entre dez caras tornará obrigatório que o comprimento da maior cadeia de caras consecutivas seja no mínimo cinco e no máximo dez.

Neste caso:

$$-\ \log_p nq = 0,90$$

não faz uma boa previsão.

Legendas das tabelas

Tabela 2. Cada linha representa o número de lançamentos de uma moeda honesta como sendo o valor de n. A previsão para o comprimento da maior cadeia de caras, através do logaritmo. E, por fim, o número de cadeias de comprimento n em que a maior cadeia de caras tem comprimento L, com L variando de zero a n.

Tabela 3. Igual a tabela 2, porém com uma linha a mais, na qual temos nove lançamentos da moeda.

Carlos J. Borges

	$-\log_p nq$	L = 0	L = 1	L = 2	L = 3	L = 4	L = 5	L = 6	L = 7	L = 8
		1								
n = 0		1								
n = 1		1	1							
n = 2		1	2	1						
n = 3	0.58	1	4	2	1					
n = 4	1.00	1	7	5	2	1				
n = 5	1.32	1	12	11	5	2	1			
n = 6	1.58	1	20	23	12	5	2	1		
n = 7	1.81	1	33	47	27	12	5	2	1	
n = 8	2.00	1	54	94	59	28	12	5	2	1

TAB 2

Carlos J. Borge

	[illegible]	L = 0	[illegible]	L = 2	[illegible]	L = 4	[illegible]	L = 6	[illegible]	L = 8	[illegible]
		1									
n = 0		1									
[illegible]		1	1								
n = 2		1	2	1							
[illegible]	0.58	[illegible]	[illegible]	[illegible]	[illegible]						
n = 4	[illegible]	1	7	5	2	1					
[illegible]	1.32	1	12	11	5	2	[illegible]				
n = 6	[illegible]	1	20	23	12	5	[illegible]	1			
[illegible]	1.81	1	33	47	27	12	[illegible]	2	1		
n = 8	[illegible]	1	54	94	69	28	[illegible]	5	2	1	
[illegible]	2.17	1	88	185	127	63	28	12	5	2	1

TAB 3

Carlos J. Borge