COMPRIMENTO MÁXIMO ESPERADO PARA AS CADEIAS DE CARAS

Carlos José Borge

Gravitation Editora: R. Dr. Luiz Migliano, 761, Ap. 54, bloco C, Cep: 05711-001, E-mail: cjborge@hotmail.com

Resumo

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 " $-\log_p nq$ "

Palavras chave: Ensaios de Bernoulli, probabilidades, passeios aleatórios, processos estocásticos, lançamentos sucessivos de moedas.

Introdução

William Feller escreveu [1]: *"O homem comum e também o filósofo K. Marbe acreditam que, depois de uma seqüência de dezessete caras, o aparecimento de uma coroa se torna mais provável. Esse argumento nada tem a ver com as imperfeições das moedas físicas; ele dá à natureza uma memória ou, na nossa terminologia, ele nega a independência estocástica entre os lançamentos sucessivos. A teoria de Marbe não pode ser refutada por argumentos lógicos, mas é rejeitada pela falta de suporte empírico."*

Quando li isso, como aluno de graduação do IFUSP, comecei a pensar: *"Por que ele escolheu o número dezessete. Será que tem algo aí?* A maioria dessas argumentações não dá em nada, mas 5%, digamos, nos fornecem algo novo sobre a natureza. Então, aproveitei que estavam saindo os primeiros computadores pessoais, programados em **basic** e pus o meu para lançar moedas. Ao observar as seqüências de caras e coroas, a lei, isto é o logaritmo apareceu rapidamente, como descrito adiante. Antes, porém, da descrição vejamos mais um dado histórico interessante".

Em 1999, Ted Hill [2], professor de Matemática do Instituto de Tecnologia da Geórgia em Atlanta, 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:

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

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 exemplos 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

$2^9 = 512 = 1 + 88 + 185 + 127 + 63 + 28 + 12 + 5 + 2 + 1$ (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.

Para 50 lançamentos com p = ½ temos:

$2^{50} = 1{,}125900 \times 10^{15}$ número total de possibilidades (espaço amostral);

$-\log_p nq = \log_2 25 = 4{,}64$ indicando Lp,máx em torno de 4 ou 5;

L = 0 corresponde a 1 possibilidade (50 coroas), com probabilidade $P(0) = 2^{-50}$;

L = 1 corresponde a $3{,}295128 \times 10^{10}$ possibilidades, com P(1) = 0,0029%;

L = 2 corresponde a $1{,}939402 \times 10^{13}$ possibilidades, com P(2) = 1,72%;

L = 3 corresponde a $1{,}748875 \times 10^{14}$ possibilidades, com P(3) = 15,53%;

L = 4 corresponde a $3{,}102290 \times 10^{14}$ possibilidades, com P(4) = 27,55%;

L = 5 corresponde a $2{,}671568 \times 10^{14}$ possibilidades, com P(5) = 23,73%;

L = 6 corresponde a 1,680508 × 10^{14} possibilidades, com P(6) = 14,93%;

L = 7 corresponde a 9,206614 × 10^{13} possibilidades, com P(7) = 8,18%;

L = 8 corresponde a 4,740258 × 10^{13} possibilidades, com P(8) = 4,21%;

L = 9 corresponde a 2,372302 × 10^{13} possibilidades, com P(9) = 2,11%;

L = 10 corresponde a 1,171618 × 10^{13} possibilidades, com P(10) = 1,04%;

.
.
.

A soma das probabilidades acima, P(1) a P(10), é igual a 99%;

Note que P(4) + P(5) = 51,28% e que P(3) + P(4) + P(5) +P(6) = 81,74%, revelando que este logaritmo é uma excelente ferramenta matemática.

GRÁFICO 1

Resultados de simulações

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.

Numa outra simulação, com computador, uma moeda foi lançada n = 1024 vezes, sendo p = 1/5 e q = 4/5 as probabilidades de cara e coroa respectivamente, e o experimento foi repetido 50 vezes.

Para os comprimentos **esperados,** de acordo com o nosso logaritmo, das maiores cadeias de caras e coroas, temos, respectivamente:

$- \log_p nq = 4{,}17$ (caras – **teórico**) e

$- \log_q np = 23{,}85$ (coroas – **teórico**)

Calculando as probabilidades P(x), com base no resultado do experimento, de as maiores cadeias terem comprimento x, obtivemos:

Os valores **medidos** para as cadeias de caras foram:

P(3) = 20%; P(4) = 58%; P(5) = 20%; P(6) = 2%, totalizando 100%. Neste experimento não foram obtidos comprimentos diferentes de 3, 4, 5 e 6.

Os valores **medidos** para as cadeias de coroas foram:

P(17) = 2% (ocorreu uma vez), P(21) = 4%, P(22) = 12%, P(23) = 10%, P(24) = 10%, P(25) = 10%, P(26) = 16%, P(27) = 6%; P(28) = 6%, P(29) = 2%, P(30) = 2%, P(31) = 8%, P(32) = 2%, P(34) = 4%, P(38) = 2%, P(42) = 2% e P(45) = 2%. Totalizando 100%.

Notar que: P(23) + P(24) = 20% e que P(22) + P(23) + P(24) + P(25) = 42%.

Casos especiais

Para p = q = ½ o logaritmo se aplica muito bem, mas quando p se afasta muito de ½, 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 para o comprimento máximo das cadeias de caras.

Já, para as cadeias de coroas a previsão é muito boa:

$$- \log_q np = 0{,}996$$

indica corretamente que a maior cadeia de coroas deverá ter comprimento em torno de um.

Se fizermos n = 110, p = 99/100 e q = 1/100, teremos um resultado análogo, isto é, uma coroa e 109 caras tem boa probabilidade de ocorrência.

A nossa previsão para o comprimento da maior cadeia de caras, será:

$$- \log_p nq = 9{,}48$$

indicando um inteiro próximo de 9 ou 10 que terá pouca chance de ocorrer. (Para uma coroa não ocorre)

Já, para as cadeias de coroas, a previsão será muito boa, isto é:

$$- \log_q np = 1{,}02$$

indicando um inteiro próximo de 1 que terá boa chance de ocorrência.

Vamos tentar entender porque o logaritmo prevê um comprimento próximo de 9 ou 10 para a maior cadeia de caras, quando na verdade esse comprimento, com boa chance, deverá ser maior que cinqüenta. (Se sair uma única coroa ele estará entre 55 e 110.)

Nesse exemplo, nq = 1,1 indica que devemos esperar, nesses 110 lançamentos, aproximadamente uma coroa, e o logaritmo enxerga isso como um único possível início de cadeia de caras. (L = 0 uma vez)

E também:

nqp = 1,089000 é maior que 1 e indica (para o logaritmo) que L = 1, espera-se que aconteça nqp vezes;

nqp^2 = 1,078110 é maior que 1 e indica (para o logaritmo) que L = 2, espera-se que aconteça nqp^2 vezes;

nqp^3 = 1,067329 é maior que 1 e indica (para o logaritmo) que L = 3, espera-se que aconteça nqp^3 vezes;

.

.

.

nqp^9 = 1,004869 é maior que 1 e indica (para o logaritmo) que L = 9, espera-se que aconteça nqp^9 vezes;

nqp^{10} = 0,994820 é **menor** que 1 e indica (para o logaritmo) que L = 10, espera-se que não chegue a acontecer.

Então, o logaritmo previu 9,48 para o comprimento da maior cadeia de caras, mas essa seqüência de valores: L = 2, L = 3, L = 4, etc. como enxerga o logaritmo, não existe fisicamente.

L deverá assumir valores variando de zero a 110, mas não como enxerga o logaritmo.

Conclusão: se p for menor ou igual a meio e se o produto nq for maior que 2 (ou provavelmente maior que 1), então o nosso logaritmo fará boas previsões para o comprimento da maior cadeia de caras.

Se p for maior que meio, e principalmente se estiver próximo de 1, então a previsão do logaritmo pode não condizer com a realidade. Principalmente se o produto nq também estiver próximo da unidade.

Este item (casos especiais) merece um estudo mais aprofundado que poderá ocorrer no futuro próximo.

Conclusão

Nesse caso, o estudo nos levou a algo novo e realmente muito interessante.

Referências

[1] William Feller, Livro: Introdução à teoria das probabilidades e suas aplicações; (Capítulo VI – Exemplos), tradução: Flávio Wagner Rodrigues e Maria Eliza Fini, São Paulo, Edigard Blücher, 1976

[2] LA RECHERCHE 316 (pp.72-75) JANVIER 1999

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.

tab50.xls.lnk

1											
1	1										
1	2	1									
1	4	2	1								
1	7	5	2	1							
1	12	11	5	2	1						
1	20	23	12	5	2	1					
1	33	47	27	12	5	2	1				
1	54	94	59	28	12	5	2	1			
1	88	185	127	63	28	12	5	2	1		
1	143	360	269	139	64	28	12	5	2	1	
1	232	694	563	303	143	64	28	12	5	2	1