A carta de Goldbach a Euler, datada de 7 de Junho de 1742, deu origem à versão moderna de sua conjectura, como atualmente difundida:

Todo inteiro par maior que 2 pode ser representado como a soma de 2 primos.

Estamos propondo uma conjectura equivalente:

Todo inteiro maior que 1 pode ser representado pela média de 2 primos.

Exemplos:

Primo $37 = (31 + 43) \div 2$;

Par $38 = (29 + 47) \div 2$;

Impar $39 = (37 + 41) \div 2$.

Então teríamos, para qualquer inteiro positivo $n > 1$, a identidade:

$$2n = p + q,$$

com p, q primos.

Sabe-se que:

$$2n = (n - k) + (n + k)$$

para k qualquer; em particular um inteiro.

E, portanto podemos ter:

$p = n - k$ e

$q = n + k$.

Desta forma, obtemos primos equidistantes de **n**, através do índex **k**, o que chamamos de **simetria** para o número **n**.

Esta simetria, envolvendo os inteiros:

$(n - k) < n$ e

$(n + k) > n$

Tem como **amplitude**:

3 ••• **n** ••• $2 \times n - 3$.

Abaixo, encontram-se várias simetrias para o número 39.

5_{-34} 7_{-32} 11_{-28} 17_{-22} 19_{-20} 31_{-8} 37_{-2} $\mathbf{39_0}$ 41_2 47_8 59_{20} 61_{22} 67_{28} 71_{32} 73_{34}

Obvio que se o próprio **n** é primo, o resultado é trivial para k = 0; porém, em nosso propósito adotamos sempre:

$\mathbf{k > 0}$,

$\mathbf{n > 3}$ e

$\mathbf{p \neq q}$.

Média aritmética simples, ordinária.

Distinguimos os inteiros ímpares dos inteiros ímpares primos.

Se o número para o qual procuramos simetria é par o índex é ímpar, e vice-versa.

Ao considerarmos a hipótese, verificamos os primeiros 2097150 inteiros consecutivos e ocorreu a confirmação do enunciado.

Mas não era o suficiente; tentamos vários outros números consecutivos (sendo sempre aleatório o primeiro deles) de maior grandeza, por exemplo:

Inteiros com 32 bits:

2326416308 ••• 2326437251

Inteiros com 64 bits:

10812083835233317544 ••• 10812083835233361798

Inteiros com 128 bits:

313545261969434692888811456477964920750 •••

313545261969434692888811456477964922750

Inteiros com 256 bits:

619232035137510808464496159340299275958970735855604059760482397121783677576 32 •••

619232035137510808464496159340299275958970735855604059760482397121783677578 00

Analisando **k**, observamos que sempre é muito pequeno em relação à **n**.

Para 2097150 inteiros, o máximo valor de k encontrado foi 1722.

Em testes aleatórios efetuados com números de 512 bits, o maior valor de k foi 70038, o que se mostrou curioso! O índex tem apenas17 bits.

A seguir o número:

131292007168910333663548776886132090673503094624508353438366940813404064932023753224857538216518806248471988525203231716334990588989835816902808492167410 69 **=**

(131292007168910333663548776886132090673503094624508353438366940813404064932023753224857538216518806248471988525203231716334990588989835816902808492167103 1) **+**

(131292007168910333663548776886132090673503094624508353438366940813404064932023753224857538216518806248471988525203231716334990588989835816902808492168111 07) **÷ 2**.

Utilizamos o algoritmo de Rabin-Miller para averiguar se os números são primos.

Tendo em vista os resultados com os primeiros inteiros consecutivos, não fomos tão rigorosos nas averiguações ulteriores e o total de iterações para o teste de primalidade foi apenas de 25 vezes para cada primo.

Mas, qual a garantia de que o resultado é sempre encontrado? E outra dúvida se impôs: como avaliar a probabilidade de encontrarmos esta simetria?

Somente para fixar uma ideia, vamos examinar o seguinte problema: temos 20 esferas perfeitas e idênticas e duas roletas ideais, uma à esquerda — **E** — e outra à direita — **D** —, cada uma com 36 células numeradas, que chamaremos de índex.

Giramos a roleta da esquerda e lançamos 11 das esferas.

Giramos a roleta da direita e lançamos as 9 esferas restantes.

Qual seria a probabilidade de se obter pelo menos uma coincidência, de forma que qualquer das 11 células ocupadas na roleta E, e qualquer das 9 células ocupadas na roleta D tivesse mesmo índex?

Apenas por conveniência, iremos investigar a questão inversa: qual seria a probabilidade $\mathfrak{P}_r$ de não se obter nenhuma coincidência? Ou seja, ao final, quando as roletas estão paradas, nenhuma das esferas tem mesmo índex!

O raciocínio: quando todas as células da roleta E estão ocupadas e lançamos a 1ª esfera na roleta D temos 36 células disponíveis. Entretanto, não queremos que seu índex coincida com nenhum dos 11 índex ocupados da outra roleta.

A probabilidade deste evento é $25 \div 36$.

Ao lançarmos a 2ª esfera já temos uma célula ocupada e, portanto, uma opção a menos, de forma que esta probabilidade é $24 \div 35$.

Sucessivamente, desta forma, as possibilidades se reduzem a cada lançamento e para a última esfera, a probabilidade é $17 \div 28$.

Para atingir o objetivo, a probabilidade de não se obter nenhuma coincidência é:

$$\mathfrak{P}_r = (25 \div 36) \times (24 \div 35) \times \bullet\bullet\bullet \times (18 \div 29) \times (17 \div 28).$$

Conhecido o resultado,

$$\mathfrak{P}_r = 0.0217,$$

Podemos, agora, responder à primeira pergunta: a probabilidade de se obter pelo menos uma coincidência é 0.9783.

Roletas com **N** células e com $\mathbf{P+Q}$ esferas, requerem equacionarmos melhor o problema, pois se os valores envolvidos são grandes o cálculo torna-se tedioso, difícil ou mesmo inexequível.

Preferimos utilizar $\mathfrak{P}_r$, a probabilidade de **não se obter nenhuma coincidência**, ao invés da probabilidade de **se obter pelo menos uma coincidência**, que se dá pelo complemento de $\mathfrak{P}_r$ para 1. E apenas $\mathfrak{P}_r$ será usada aqui!

Inexequível: parece simples distinguir números pares de ímpares! Basta ver o algarismo das unidades. Contudo, não é imediato para um número da ordem de um googol **escrito** na base 5! Observe que 10_5 não é divisível por 2.

E se temos um googolplex? Vide Kasner & Newman.

Formulando, temos:

$$\mathfrak{P}_r = [(N-P) \div N] \times [(N-P-1) \div (N-1)] \times \bullet\bullet\bullet$$
$$\bullet\bullet\bullet \times [(N-P-Q+1) \div (N-Q+1)].$$

E a seguinte identidade combinatória, para inteiros a > b > m > 0, é útil:

$$\mathbb{C}\{^{b}{}_{m}\} \div \mathbb{C}\{^{a}{}_{m}\} = \{[b! \div (b-m)!]\} \div \{[a! \div (a-m)!]\}$$
$$= [(b \div a)] \times [(b-1) \div (a-1)] \times \bullet\bullet\bullet$$
$$\bullet\bullet\bullet \times [(b-m+1) \div (a-m+1)].$$

Assim, com nossas variáveis, se N > P > Q e $N - P \geq Q$, temos:

$$\mathfrak{P}_r = \mathbb{C}\{^{N-P}{}_{Q}\} \div \mathbb{C}\{^{N}{}_{Q}\}.$$

E, para ilustrar, no caso das roletas teríamos: $\mathbb{C}\{^{25}{}_{9}\} \div \mathbb{C}\{^{36}{}_{9}\}$.

Voltando à nossa conjectura, vamos investigar o que ocorre com primos distribuídos entre inteiros — usando o mesmo modelo anterior — fixando certo número **n** e considerando a amplitude de **N** inteiros:

Menores que **n** contendo **P** primos e

Maiores que **n** contendo **Q** primos.

Já vimos como calcular a probabilidade de não encontrarmos nenhum par de esferas sob índex equivalente e temos uma questão análoga, sendo os primos:

$p = n - k$ e

$q = n + k$.

Se N, e, portanto P e Q, são de grandezas elevadas é difícil obter a probabilidade como fizemos, pois mesmo sendo N conhecido, como saber o valor de P e Q?

Primeiramente, podemos usar um artificio!

Não é difícil verificar que:

$$[(N-P) \div N] > [(N-P-1) \div (N-1)] > \bullet\bullet\bullet$$
$$\bullet\bullet\bullet > [(N-P-Q+1) \div (N-Q+1)].$$

E, consequentemente, podemos fazer:

$$\mathfrak{P}_r = [(N-P) \div N]^{Q}$$

tendo em vista que este valor é maior que $\mathbb{C}\{^{N-P}{}_{Q}\} \div \mathbb{C}\{^{N}{}_{Q}\}$.

A seguir, sabe-se o que o Teorema dos Números Primos (TNP) nos assegura:

$$\Pi(x) \approx x \div \log(x).$$

Para o fatorial de números muito grandes é melhor usar a aproximação de Stirling.

TNP: O teorema descreve a distribuição dos números primos entre inteiros e foi demonstrado independentemente por Jacques Hadamard e Charles Jean de la Vallée-Poussin em 1896, através do estudo da função ζ de Bernhard Riemann. O teorema nos assegura que a quantidade de primos menores (ou eventualmente igual) que x é proporcional à razão entre x e o $\log_e(x)$.

O que nos permite dizer, dentro da amplitude disponível, com P ≥ Q:

$$P \approx N \div \log(N) \qquad \textbf{primos} < \textbf{n}$$

$$Q \approx 2N \div \log(2N) - N \div \log(N) \qquad \textbf{primos} > \textbf{n}$$

Temos:

$$\Re_r = [1 - (P \div N)]^{Q}$$

E substituindo P e Q, temos:

$$\Re_r = \{1 - [1 \div \log(N)]\}^{[2N \div \log(2N) - N \div \log(N)]}.$$

Para a conjectura ser válida precisávamos, por fim, demonstrar que $\Re_r$ tende a zero quando **n** tende a infinito e... então nos encontramos diante de um aparente paradoxo: o intuitivo limite da função não era zero e os cálculos indicavam que sim.

Mas, prosseguimos e, malgrado o tempo que demandamos para calcular este limite, não conseguimos obtê-lo; todas tentativas foram infrutíferas.

Entretanto, recorrendo à Internet vimos que o problema já é consagrado entre acadêmicos e, de fato, o limite da função é zero.

Assim foi que obtivemos o resultado esperado, e desta forma,

A probabilidade de não se obter nenhuma coincidência é:

$$\textbf{Lim}_{N \to \infty} \Re_r = \textbf{0}$$

Ou seja, provavelmente, sempre haverá pelo menos uma coincidência.

Avaliar os primeiros números foi fácil e muito mais podemos obter simplesmente computando exaustivamente através de iterações continuas.

Porem, em vista do que temos: um cálculo probabilístico; poderia haver algum **n** para o qual a conjectura falhe, entre o último que venha a ser obtido e o infinito?

Ivan Gondim Leichsenring
Apex Algoritmos [apex.eti.br]
ivan@apex.eti.br

Se você puder ajudar a tornar este texto mais legível, eu agradeço.

Intuitivo... para nós!

São várias organizações que dispõem o resultado deste limite, entre elas:
http://www.MathPortal.org http:// www.DerivativeCalculator.net http:// www.WolframAlpha.com