Harmonic Number

1 Introduction

The term *Harmonic number* has multiple meanings.

In mathematics, the n-th harmonic number is the sum of the reciprocals of the first n natural numbers:

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n} = \sum_{i=1}^{n} \frac{1}{n}$$

This also equals n times the inverse of the *harmonic mean* of these natural numbers.

Harmonic numbers were studied in antiquity and are important in various branches of number theory. They are sometimes loosely termed harmonic series, are closely related to the Riemann zeta function, and appear in the expressions of various special functions.

The associated harmonic series grows without limit, albeit very slowly, roughly approaching the natural logarithm function.

In 1737, Leonhard Euler used the divergence of this series to provide a new proof of the infinity of prime numbers. His work was extended into the complex plane by Bernhard Riemann in 1859, leading directly to the celebrated Riemann hypothesis about the distribution of prime numbers.

Figure 1: The harmonic number $H_{n,1}$ with $n = \lfloor x \rfloor$ (red line) with its asymptotic limit $\gamma + \ln(x)$ (blue line).

When the value of a large quantity of items has a Zipf's law distribution, the total value of the n most-valuable items is the n-th harmonic number. This leads to a variety of surprising conclusions in the Long Tail and the theory of network value.

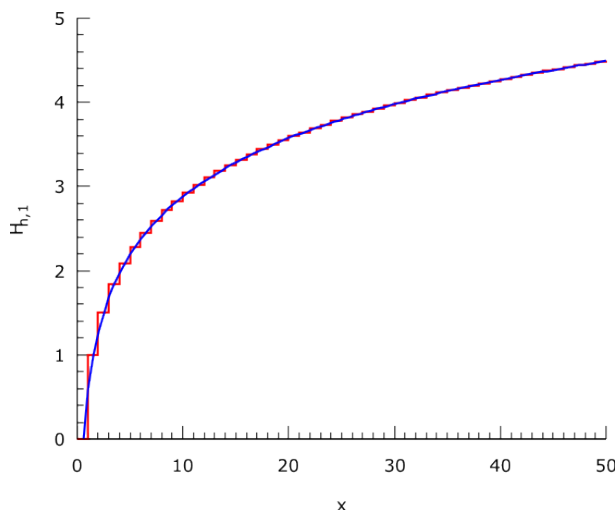Bertrand's postulate entails that, except for the case n=1, the harmonic numbers are never integers.

2   Identities involving harmonic numbers

By definition, the harmonic numbers satisfy the *recurrence relation*

$$H_n = H_{n-1} + \frac{1}{n}$$

They also satisfy the series Identity

$$\sum_{k=1}^{n} H_k = (n+1)H_n - n$$

The Harmonic numbers are connected the *Stirling Numbers* of first kind

$$H_n = \frac{1}{n!}\binom{n+1}{2}$$

The functions

$$f_n(x) = \frac{x^n}{n!}(\log x - H_n)$$

satisfy the property

$$f'_n(x) = f_{n-1}(x)$$

In particular

$$f_1(x) = x(\log x - 1)$$

3   Eulerian Representation

An integral representation is given by

$$H_n = \int_0^1 \frac{1 - x^n}{1 - x} dx$$

The equality above is obvious by simple algebraic identity

$$\frac{1 - x^n}{1 - x} = 1 + x + ... + x^{n-1}$$

Using a simple integral transform $x = 1 - u$, an elegant combinatorial expression for H_n is

$$
\begin{aligned}
H_n &= \int_0^1 \frac{1 - x^n}{1 - x} dx \\
&= -\int_1^0 \frac{1 - (1 - u)^n}{u} dx \\
&= \int_0^1 \frac{1 - (1 - u)^n}{u} dx \\
&= \int_0^1 \left[\sum_{k=1}^n (-1)^{k-1} \binom{n}{k} u^{k-1} \right] dx \\
&= \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} \int_0^1 u^{k-1} dx \\
&= \sum_{k=1}^n (-1)^{k-1} \frac{1}{k} \binom{n}{k}
\end{aligned}
$$

The same representation can be produced by using the third *Retkes Identity* by setting $x_1 = 1, ..., x_n = n$ and using the fact that

$$\Pi_k(1, ..., n) = (-1)^{n-k}(k-1)!(n-k)!$$

$$H_n = H_{n,1} = \sum_{k=1}^n \frac{1}{k} = (-1)^{n-1} n! \sum_{k=1}^n \frac{1}{k^2 \Pi_k(1, ..., n)} = \sum_{k=1}^n (-1)^{k-1} \frac{1}{k} \binom{n}{k}$$

4   Approximation of H_n by $\int_0^1 \frac{1}{x} dx$

The harmonic sum is the sum of reciprocals of the positive integers.
We have

$$\frac{1}{n} < \int_{n-1}^n \frac{dt}{t}, n \ge 2$$

$$\frac{1}{n} > \int_n^{n+1} \frac{dt}{t}, n \ge 1$$

Hence we can say, by a summation of n upto N

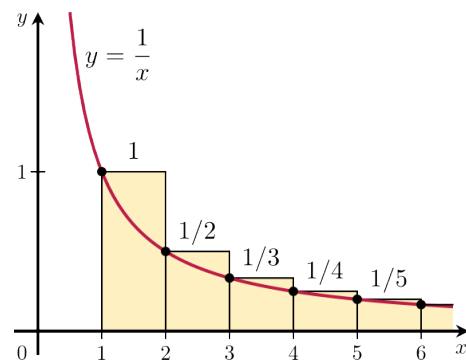$$\int_1^{N+1} \frac{dt}{t} < \sum_{n=1}^N \frac{1}{n} < 1 + \int_1^N \frac{dt}{t}$$

Figure 2: Graph demonstrating a connection between harmonic numbers and the natural logarithm. The harmonic number H_n can be interpreted as a Riemann sum of the integral: $\int_1^{n+1} \frac{1}{x} dx = \ln(n + 1)$

3

$$\log\left(N+1\right) < \sum_{n=1}^{N} \frac{1}{n} < 1 + \log N$$

since $\log N < \log\left(N+1\right)$ we can write

$$\log N < \sum_{n=1}^{N} \frac{1}{n} < 1 + \log N$$

Now we can further write

$$\sum_{n=1}^{N} \frac{1}{n} = \log N + O(1)$$

Hence we can introduce some constant $\lambda < 1$ such that

$$\left| \sum_{n=1}^{N} \frac{1}{n} - \log N \right| \leq \lambda, \forall N$$

The nth harmonic number is about as large as the *natural logarithm* of n. The reason is that the sum is approxiamted by the integral

$$\int_{1}^{n} \frac{1}{x} dx$$

whose value is *ln(n)*

Can we get any better?This is all wonderful, but can we do better? That is, can we be more precise about the error.
Lets see, begin with

$$a_n = \frac{1}{n} - \int_{n}^{n+1} \frac{dt}{t}$$

Integrating it and using properties of logarithm

$$a_n = \frac{1}{n} - log(1 + \frac{1}{n}) \tag{1}$$

and expanding via the Taylor series for $\log\left(1+x\right)$ which converges for $-1 \leq x \leq 1$, we get

$$a_n = \frac{1}{n} - \left(\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3}... \right) = \frac{1}{2n^2} - \frac{1}{3n^3} \tag{2}$$

This is an alternating series where terms decrease in absolute value, so we see that

$$0 < a_n < \frac{1}{2n^2} \tag{3}$$

A consequence of this inequality is that the sum

$$\sum_{n=1}^{\infty} a_n \tag{4}$$

4

is a constant. Call this constant γ. Now we can say that

The values of the sequence $H_n - ln(n)$ decreases monotonically towards the *limit*

$$\lim_{n \to \infty} (H_n - ln(n)) = \gamma$$

where $\gamma \sim 0.5772156649$ is the *Euler-Mascheroni Constant*. The correspondng *asymptotic expansion* as $n \to \infty$ is

$$H_n \sim ln\,n + \gamma + \frac{1}{2n} - \sum_{k=1}^{\infty} \frac{B_{2k}}{skn^{2k}} = ln\,n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} - \cdots$$

where B_k are the *Bernoulli Numbers*.

References