

## On the Means of the Values of Prime Counting Function

Mehdi Hassani

Department of Mathematics, University of Zanjan, University Blvd.,  
 45371-38791, Zanjan, Iran.

E-mail: mehdi.hassani@znu.ac.ir

**ABSTRACT.** In this paper, we investigate the means of the values of primes counting function  $\pi(x)$ . First, we compute the arithmetic, the geometric, and the harmonic means of the values of this function, and then we study the limit value of their ratio.

**Keywords:** Prime number, Prime counting function, Means of the values of function.

**2000 Mathematics subject classification:** 11N05, 26E60.

### 1. INTRODUCTION AND SUMMARY OF THE RESULTS

**1.1. Means of the values of primes counting function.** Assume that  $(a_n)_{n \in \mathbb{N}}$  is a strictly positive real sequence. The arithmetic mean of the numbers  $a_1, a_2, \dots, a_n$  is defined by

$$A(a_1, \dots, a_n) = \frac{1}{n} \sum_{k=1}^n a_k.$$

The geometric and harmonic means of the these numbers, defined in terms of arithmetic mean, respectively, by

$$G(a_1, \dots, a_n) = e^{A(\log a_1, \dots, \log a_n)},$$

and

$$H(a_1, \dots, a_n) = \frac{1}{A(\frac{1}{a_1}, \dots, \frac{1}{a_n})}.$$

---

Received 15 March 2014; Accepted 13 February 2018

©2018 Academic Center for Education, Culture and Research TMU

All of the above means are special cases of the so-called generalized mean with parameter  $r \in \mathbb{R}$ , defined by

$$M_r(a_1, \dots, a_n) = (A(a_1^r, \dots, a_n^r))^{\frac{1}{r}}.$$

We note that  $M_1 = A$ ,  $M_0 = \lim_{r \rightarrow 0} M_r = G$ , and  $M_{-1} = H$ .

Analogue to the above discrete case, we assume that for some fixed  $a \in \mathbb{R}$  the functions  $f$  with  $f : [a, \infty) \rightarrow (0, \infty)$  is an integrable function. For any real number  $b > 0$ , we define the arithmetic, the geometric and the harmonic means of the values of  $f$  over the interval  $[a, b+a]$  respectively by

$$A_b(f) = \frac{1}{b} \int_a^{b+a} f(t) dt, \quad G_b(f) = e^{A_b(\log f)}, \quad \text{and} \quad H_b(f) = \frac{1}{A_b(\frac{1}{f})}.$$

More generally, we define the generalized mean with parameter  $r \in \mathbb{R}$  by

$$M_{b,r}(f) = A_b(f^r)^{\frac{1}{r}}.$$

Our intention in writing this paper is to investigate means of the values of primes counting function  $\pi(x)$ , which denotes the number of primes not exceeding  $x$ . Since  $\pi(t) = 0$  for  $t < 2$ , and  $\pi(t) > 0$  for  $t \geq 2$ , we take the mean values of this function over the interval  $[2, b+2]$ . We prove the following.

**Theorem 1.1.** *Assume that  $A_b(\pi)$ ,  $G_b(\pi)$ , and  $H_b(\pi)$  denote the arithmetic, the geometric and the harmonic means of the values of the prime counting function  $\pi(x)$ , over the interval  $[2, b+2]$  with  $b > 5$ , and  $p_n$  denotes the largest prime not exceeding  $b+2$ . Then, as  $n \rightarrow \infty$  (and equivalently  $b \rightarrow \infty$ ), we have*

$$A_b(\pi) = \frac{n}{2} + O\left(\frac{\log n}{n}\right), \quad (1.1)$$

$$G_b(\pi) = e^{\log n + O(1)}, \quad (1.2)$$

and

$$H_b(\pi) = \frac{2n}{\log \log n} \left(1 + O\left(\frac{1}{\log \log n}\right)\right). \quad (1.3)$$

To prove the above theorem, we need to compute  $\int_2^{b+2} g(\pi(t)) dt$  for  $g(x) = x$ ,  $g(x) = \log x$ , and  $g(x) = \frac{1}{x}$ . In Section 2 we give a result, which enables us to compute the above mentioned integral for a certain function  $g$ , covering the required cases.

**1.2. The ratio of the arithmetic and geometric means.** For the sequence consisting of positive integers, Stirling's approximation for  $n!$  implies that

$$\frac{A(1, \dots, n)}{G(1, \dots, n)} = \frac{e}{2} + O\left(\frac{\log n}{n}\right). \quad (1.4)$$

Motivated by this fact, recently we obtained similar asymptotic result concerning the sequence of prime numbers, by proving

$$\frac{A(p_1, \dots, p_n)}{G(p_1, \dots, p_n)} = \frac{e}{2} + O\left(\frac{1}{\log n}\right), \quad (1.5)$$

where as usual  $p_n$  denotes the  $n$ th prime number (see [2]).

Similar to the above, we denote

$$\frac{A}{G}(f) = \lim_{b \rightarrow \infty} \frac{A_b(f)}{G_b(f)},$$

provided the above limit exists. For instance, if we let  $f(x) = [x]$ , the integer part of real  $x$ , then over the interval  $[1, b+1]$  we have

$$A_b(f) = \frac{1}{n} \int_1^{n+1} [t] dt = \frac{1}{n} \sum_{k=1}^n \int_k^{k+1} [t] dt = \frac{1}{n} \sum_{k=1}^n k = A(1, 2, \dots, n),$$

and  $G_b(f) = G(1, 2, \dots, n)$ , which gives the limit relation (1.4) for  $\frac{A}{G}(f)$ . Moreover, analogously to (1.4), one may consider  $\frac{A}{G}(f)$  for  $f(x) = x$ . For the case of prime numbers, the prime number theorem asserts that  $p_n \sim n \log n$  as  $n \rightarrow \infty$ . Thus, analogously to the limit relation (1.5), one may consider  $\frac{A}{G}(f)$  for  $f(x) = x \log x$ . Straightforward computations imply that  $\frac{A}{G}(f) = \frac{e}{2}$  for  $f(x) = x$  and  $f(x) = x \log x$ . We note that the appearance of the similar limit value  $\frac{e}{2}$  is not a global property. For example, a similar computation as the above implies that  $\frac{A}{G}(f) = 1$  for  $f(x) = \log x$ . In general,  $A_b(f) \geq G_b(f)$ , and we observe that the limit value of the ratio  $\frac{A}{G}$  could be any arbitrary real number  $\beta \geq 1$ , as the following constructive result confirms.

**Theorem 1.2.** *For any real number  $\beta \geq 1$  there exists a real positive function  $f$  such that*

$$\frac{A}{G}(f) = \beta.$$

*Remark 1.3.* One may ask about existence and the value of  $\lim_{b \rightarrow \infty} \frac{A_b(f)}{G_b(f)}$ , for  $f(x) = \pi(x)$ . The prime number theorem asserts that  $\pi(x) \sim \frac{x}{\log x}$ , as  $x \rightarrow \infty$ . For the function  $f(x) = \frac{x}{\log x}$ , straightforward computation implies that  $\frac{A}{G}(f) = \frac{e}{2}$ . But, our computations in (1.1) and (1.2), mainly those of geometric mean values, is not enough strong to get similar result for  $\pi(x)$ . Our argument in the next section, supports that the value of  $\lim_{b \rightarrow \infty} \frac{A_b(f)}{G_b(f)}$  for  $f(x) = \pi(x)$ , if exists, is closely related to the value of the limit

$$\lim_{n \rightarrow \infty} \frac{S(n) - \frac{1}{2}np_n}{n^2}, \quad (1.6)$$

provided it exists, where  $S(n) = \sum_{k=1}^n p_k$ . In [2] we prove that

$$\frac{n}{2}p_n - \frac{9}{4}n^2 < S(n) < \frac{n}{2}p_n - \frac{1}{12}n^2,$$

where the left hand side inequality is valid for any integer  $n \geq 2$ , and the right hand side inequality is valid for any integer  $n \geq 10$ . Thus, the value of the limit (1.6) lies in the interval  $[-\frac{9}{4}, -\frac{1}{12}]$ . We guess that its true value is  $-\frac{1}{4}$ , and consequently, we conjecture that the true value of  $O(1)$  in (1.2) is also  $-\frac{1}{4}$ , and hence,  $\frac{A}{G}(f) = \frac{\sqrt[4]{6}}{2}$  for  $f(x) = \pi(x)$ .

## 2. AN AUXILIARY GENERAL RESULT

The following results prepare the main tool of explicit and approximate computing several means of the values of  $\pi(x)$ .

**Lemma 2.1.** *For  $S(n) = \sum_{k=1}^n p_k$  and  $g$  be continuously differentiable on  $[1, n-1]$ , we have*

$$\begin{aligned} I &:= \int_e^{n-1} S([t]+1)(g'(t+1) - g'(t)) \, dt \\ &= S(n)(g(n) - g(n-1)) + 2g(1) - c_g - \sum_{k=1}^{n-1} (g(k+1) - g(k))p_{k+1}, \end{aligned}$$

where  $c_g$  is a constant defined in terms of  $g$ .

*Proof.* We let  $I = \int_1^{n-1} - \int_1^e := I_3 - \int_1^e$  with

$$\begin{aligned} I_3 &:= \int_1^{n-1} S([t]+1)(g'(t+1) - g'(t)) \, dt \\ &= \sum_{k=1}^{n-2} \int_k^{k+1} S(k+1)(g'(t+1) - g'(t)) \, dt \\ &= \sum_{k=1}^{n-2} S(k+1)(g(k+2) - g(k+1)) - \sum_{k=1}^{n-2} S(k+1)(g(k+1) - g(k)) \\ &= \sum_{k=2}^{n-1} S(k)(g(k+1) - g(k)) - \sum_{k=1}^{n-2} S(k+1)(g(k+1) - g(k)) \\ &= S(n)(g(n) - g(n-1)) - 2g(2) + 2g(1) - \sum_{k=1}^{n-1} p_{k+1}(g(k+1) - g(k)). \end{aligned}$$

This completes the proof.  $\square$

**Theorem 2.2.** *Assume that  $b > 0$  is a real number, and  $p_n$  denotes the largest prime not exceeding  $b+2$ . Also, assume that  $g : (0, +\infty) \rightarrow \mathbb{R}$  is a continuous function. Then, we have*

$$\begin{aligned} \int_2^{b+2} g(\pi(t)) \, dt &= g(n)(b+2 - p_n) + \sum_{k=1}^{n-1} (p_{k+1} - p_k)g(k) \\ &= g(n)(b+2) - 2g(1) - \sum_{k=1}^{n-1} (g(k+1) - g(k))p_{k+1}. \end{aligned} \quad (2.1)$$

Moreover, if  $g$  is continuously differentiable on the interval  $[1, n-1]$  and  $g'(t) = \frac{d}{dt}g(t)$ , then for any  $b > 5$  we have

$$\int_2^{b+2} g(\pi(t)) dt = (b+2)g(n) - S(n)(g(n) - g(n-1)) \quad (2.2)$$

$$+ c_g + \int_e^{n-1} S([t] + 1)\Delta(t) dt,$$

where  $S(n) = \sum_{k=1}^n p_k$ ,  $c_g = 10g(e+1) - 10g(e) - 5g(3) + 2g(2) + g(1)$ , and  $\Delta(t) := g'(t+1) - g'(t)$ . Also, as  $n \rightarrow \infty$  (and equivalently  $b \rightarrow \infty$ ), we have

$$\int_2^{b+2} g(\pi(t)) dt = G(n) + O(R(n)), \quad (2.3)$$

where

$$G(n) = \left( g(n) - \frac{n}{2}(g(n) - g(n-1)) \right) n\ell(n) + c_g + \frac{1}{2} \int_e^{n-1} t^2 \ell(t) \Delta(t) dt,$$

with  $\ell(t) = \log t + \log \log t$ , and

$$R(n) = \left( g(n) + n(g(n) - g(n-1)) \right) n + \int_e^{n-1} t^2 \Delta(t) dt.$$

As more as, we have

$$\frac{1}{b} \int_2^{b+2} g(\pi(t)) dt = \frac{1}{2n\ell(n)} \int_e^{n-1} t^2 \ell(t) \Delta(t) dt + \frac{c_g}{n\ell(n)} \quad (2.4)$$

$$+ \left( g(n) - \frac{n}{2}(g(n) - g(n-1)) \right) + O\left( \frac{\frac{G(n)}{\log n} + R(n)}{n \log n} \right).$$

*Proof.* Since  $p_n$  is the largest prime not exceeding  $b+2$ , one may write

$$\int_2^{b+2} g(\pi(t)) dt = \int_2^{p_n} g(\pi(t)) dt + \int_{p_n}^{b+2} g(\pi(t)) dt := I_1 + I_2,$$

say, respectively. We note that  $\pi(t) = k-1$  if and only if  $p_{k-1} \leq t < p_k$ . Thus, we obtain  $I_2 = g(n)(b+2 - p_n)$ , and

$$I_1 = \sum_{k=2}^n \int_{p_{k-1}}^{p_k^-} g(\pi(t)) dt = \sum_{k=2}^n g(k-1)(p_k - p_{k-1}) := T_g(n-1),$$

say. This implies validity of (2.1). Now, we apply the truth of Lemma 2.1 to (2.2). Note that we take  $b > 5$  to guarantee  $n \geq 4$ . Finally, we deduce (2.3) by applying the known approximations (see [2] and [1], respectively)

$$S(n) = \frac{1}{2}np_n + O(n^2), \quad \text{as } n \rightarrow \infty, \quad (2.5)$$

and

$$p_n = n(\ell(n) + O(1)), \quad \text{as } n \rightarrow \infty, \quad (2.6)$$

from which we get  $S([t] + 1) = \frac{t^2}{2}\ell(t) + O(t^2)$ , and so

$$\int_e^{n-1} S([t] + 1)\Delta(t) dt = \frac{1}{2} \int_e^{n-1} t^2 \ell(t) \Delta(t) dt + O\left(\int_e^{n-1} t^2 \Delta(t) dt\right).$$

Moreover, the relations (2.5) and (2.6) yield

$$S(n) = \frac{1}{2}n^2\ell(n) + O(n^2).$$

Also, we have  $p_n \leq b + 2 \leq p_{n+1}$ , from which by applying (2.6) we get

$$b + 2 = n(\ell(n) + O(1)).$$

By applying the three last relations in (2.2), we obtain validity of (2.3). Also, we use  $b = n(\ell(n) + O(1))$  to get

$$\frac{1}{b} = \frac{1}{n\ell(n)} \left(1 + O\left(\frac{1}{\log n}\right)\right).$$

This implies validity of (2.4), and completes the proof.  $\square$

*Remark 2.3.* The constants of  $O$ -terms in the relations (2.5) and (2.6) are known explicitly (see [2] and [3]). Thus, one may compute the constants of  $O$ -terms in the relations (2.3) and (2.4) for the given function  $g$ .

### 3. PROOFS OF THE OTHER RESULTS

We will need some integration formulas, recalled here briefly. We recall that  $\text{Li}$  is the logarithmic integral function defined by

$$\text{Li}(x) = \int_0^x \frac{1}{\log t} dt,$$

where we take the Cauchy principal value of the integral. Integration by parts implies that

$$\text{Li}(x) = \frac{x}{\log x} \sum_{k=0}^m \frac{k!}{\log^k x} + O\left(\frac{x}{\log^{m+2} x}\right), \quad (3.1)$$

for any integer  $m \geq 0$ . A simple computation verifies that

$$\int \log \log x dx = x \log \log x - \text{Li}(x), \quad (3.2)$$

and this gives

$$\int \ell(x) dx = \int \log(x \log x) dx = x \log x + x \log \log x - x - \text{Li}(x). \quad (3.3)$$

Moreover, by elementary computations, we have

$$\int \frac{\ell(x)}{x} dx = \frac{1}{2} \log^2 x + \log x \log \log x - \log x. \quad (3.4)$$

*Proof of Theorem 1.1.* We utilize the statement of Theorem 2.2 with  $g(x) = x$ . We have  $c_g = 0$ , and  $\Delta(t) = 0$ . Thus, we get  $G(n) = \frac{1}{2}n^2\ell(n)$ , and  $R(n) = 2n^2$ , and these imply (1.1).

To compute the geometric mean, we apply the statement of Theorem 2.2 with  $g(x) = \log x$ . We have

$$\Delta(t) = \frac{1}{t^2} \left( -1 + \frac{1}{t} - \frac{1}{t(t+1)} \right).$$

Hence, we obtain

$$\int_e^{n-1} t^2 \Delta(t) dt = -n + \log n + e + 1 - \log(e+1) = O(n),$$

and

$$t^2 \ell(t) \Delta(t) = -\ell(t) + \frac{\ell(t)}{t} - \frac{\ell(t)}{t(t+1)},$$

from which by using the relations (3.3) and (3.4), together with the relation (3.1), we deduce that

$$\int_e^{n-1} t^2 \ell(t) \Delta(t) dt = -n\ell(n) + O(n).$$

Also, (with  $g(x) = \log x$ ) we have

$$g(n) - \frac{n}{2}(g(n) - g(n-1)) = \log n - \frac{1}{2} + O\left(\frac{1}{n}\right),$$

and

$$g(n) + n(g(n) - g(n-1)) = \log n + 1 + O\left(\frac{1}{n}\right).$$

Therefore  $G(n) = \ell(n)(n \log n - n) + O(n)$ , and  $R(n) = n \log n + O(n)$ . Thus, we obtain

$$\frac{1}{b} \int_2^{b+2} \log \pi(t) dt = \log n + O(1),$$

and this gives (1.2).

Similarly, we compute the harmonic mean, by using Theorem 2.2 with  $g(x) = \frac{1}{x}$ . We have

$$\Delta(t) = \frac{2t+1}{(t(t+1))^2} = \frac{2}{t^3} + O\left(\frac{1}{t^4}\right).$$

Thus,  $\int_e^{n-1} t^2 \Delta(t) dt = O(\log n)$ , and  $\int_e^{n-1} t^2 \ell(t) \Delta(t) dt = \log^2 n + 2 \log n \log \log n + O(\log n)$ . Also, (with  $g(x) = \frac{1}{x}$ ) we have  $g(n) - \frac{n}{2}(g(n) - g(n-1)) = O\left(\frac{1}{n}\right)$  and  $g(n) + n(g(n) - g(n-1)) = O\left(\frac{1}{n}\right)$ . So,  $G(n) = \frac{1}{2} \log^2 n + \log n \log \log n + O(\log n)$ , and  $R(n) = O(\log n)$ . By using the expansion

$$\frac{1}{\ell(n)} = \frac{\log \log n}{\log^2 n} \left( 1 + O\left(\frac{\log \log n}{\log n}\right) \right),$$

which is valid as  $n \rightarrow \infty$ , we obtain

$$\frac{1}{b} \int_2^{b+2} \frac{1}{\pi(t)} dt = \frac{\log \log n}{2n} + O\left(\frac{1}{n}\right).$$

and this gives (1.3). The proof is completed.  $\square$

*Proof of Theorem 1.2.* For any real number  $\eta \geq 0$ , we set  $f(x) = x^\eta$ . We have

$$A_b(f) = \frac{(b+1)^{\eta+1} - 1}{b(\eta+1)}, \quad \text{and} \quad G_b(f) = \exp \left( \eta \left( \frac{b+1}{b} \log(b+1) - 1 \right) \right).$$

Therefore, we obtain

$$\frac{A}{G}(f) = \frac{e^\eta}{\eta+1} := v(\eta),$$

say. We note that  $\frac{d}{d\eta} v(\eta) = v(\eta) \frac{\eta}{\eta+1}$ , hence  $v(\eta)$  is strictly increasing for  $\eta \geq 0$ , as well as  $v(0) = 1$  and  $\lim_{\eta \rightarrow \infty} v(\eta) = \infty$ . Thus, for any real number  $\beta \geq 1$  there exists a real number  $\eta \geq 0$  such that  $v(\eta) = \beta$ , as desired.  $\square$

#### ACKNOWLEDGMENTS

I express my gratitude to the anonymous referee(s) for careful reading of the manuscript and giving the many valuable suggestions and corrections, which improved the presentation of the paper.

#### REFERENCES

1. M. Cipolla, La determinazione assintotica dell'nimo numero primo (asymptotic determination of the  $n$ -th prime), *Rendiconto dell'Accademia delle Scienze Fisiche e Matematiche*, Series 3, **8**, (1902), 132–166.
2. M. Hassani, On the ratio of the arithmetic and geometric means of the prime numbers and the number  $e$ , *International Journal of Number Theory*, **9**(6), (2013), 1593–1603.
3. J. B. Rosser, L. Schoenfeld, Approximate formulas for some functions of prime numbers, *Illinois Journal of Mathematics*, **6**, (1962), 64–94.