# 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 $a_1, a_2, \ldots, a_n$ is defined by

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

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

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

and

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

---

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

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

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

Analogue to the above discrete case, we assume that for some fixed $a \in \mathbb{R}$ the functions $f$ with $f : [a, \infty) \to (0, \infty)$ 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

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

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 $t \geqslant 2$, 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 $n \to \infty$ (and equivalently $b \to \infty$), we have*

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

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

*and*

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

To prove the above theorem, we need to compute $\int_2^{b+2} g(\pi(t)) \, \mathrm{d}t$ 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

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

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

$$\frac{A(p_1,\ldots,p_n)}{G(p_1,\ldots,p_n)} = \frac{\mathrm{e}}{2} + O\Big(\frac{1}{\log n}\Big), \tag{1.5}$$

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

Similar to the above, we denote

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

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

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

and $G_b(f) = G(1, 2, \ldots, n)$, 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 $n \to \infty$. 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 $\frac{A}{G}(f) = \frac{\mathrm{e}}{2}$ for $f(x) = x$ and $f(x) = x\log x$. We note that the appearance of the similar limit value $\frac{\mathrm{e}}{2}$ 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, $A_b(f) \geqslant G_b(f)$, and we observe that the limit value of the ratio $\frac{A}{G}$ could be any arbitrary real number $\beta \geqslant 1$, as the following constructive result confirms.

**Theorem 1.2.** *For any real number $\beta \geqslant 1$ 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 $\lim_{b\to\infty}\frac{A_b(f)}{G_b(f)}$, for $f(x) = \pi(x)$. The prime number theorem asserts that $\pi(x) \sim \frac{x}{\log x}$, as $x \to \infty$. For the function $f(x) = \frac{x}{\log x}$, straightforward computation implies that $\frac{A}{G}(f) = \frac{\mathrm{e}}{2}$. 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 $\lim_{b\to\infty}\frac{A_b(f)}{G_b(f)}$ for $f(x) = \pi(x)$, if exists, is closely related to the value of the limit

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

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 $n \geqslant 2$, and the right hand side inequality is valid for any integer $n \geqslant 10$. 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, $\frac{A}{G}(f) = \frac{\sqrt[4]{e}}{2}$ 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*

$$I := \int_{e}^{n-1} S([t]+1)(g'(t+1) - g'(t))\, \mathrm{d}t$$

$$= S(n)(g(n) - g(n-1)) + 2g(1) - c_g - \sum_{k=1}^{n-1}(g(k+1) - g(k))p_{k+1},$$

*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

$$I_3 := \int_1^{n-1} S([t]+1)(g'(t+1) - g'(t))\, \mathrm{d}t$$

$$= \sum_{k=1}^{n-2} \int_k^{k+1} S(k+1)(g'(t+1) - g'(t))\, \mathrm{d}t$$

$$= \sum_{k=1}^{n-2} S(k+1)(g(k+2) - g(k+1)) - \sum_{k=1}^{n-2} S(k+1)(g(k+1) - g(k))$$

$$= \sum_{k=2}^{n-1} S(k)(g(k+1) - g(k)) - \sum_{k=1}^{n-2} S(k+1)(g(k+1) - g(k))$$

$$= S(n)(g(n) - g(n-1)) - 2g(2) + 2g(1) - \sum_{k=1}^{n-1} p_{k+1}(g(k+1) - g(k)).$$

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 $g : (0, +\infty) \to \mathbb{R}$ is a continuous function. Then, we have*

$$\int_2^{b+2} g(\pi(t))\, \mathrm{d}t = g(n)(b+2-p_n) + \sum_{k=1}^{n-1}(p_{k+1} - p_k)g(k) \qquad (2.1)$$

$$= g(n)(b+2) - 2g(1) - \sum_{k=1}^{n-1}(g(k+1) - g(k))p_{k+1}.$$

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

$$\int_2^{b+2} g(\pi(t))\,\mathrm{d}t = (b+2)g(n) - S(n)(g(n) - g(n-1)) \tag{2.2}$$

$$+ c_g + \int_{\mathrm{e}}^{n-1} S([t]+1)\Delta(t)\,\mathrm{d}t,$$

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

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

*where*

$$G(n) = \left(g(n) - \frac{n}{2}(g(n) - g(n-1))\right) n\ell(n) + c_g + \frac{1}{2}\int_{\mathrm{e}}^{n-1} t^2 \ell(t)\Delta(t)\,\mathrm{d}t,$$

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

$$R(n) = \left(g(n) + n(g(n) - g(n-1))\right)n + \int_{\mathrm{e}}^{n-1} t^2 \Delta(t)\,\mathrm{d}t.$$

*As more as, we have*

$$\frac{1}{b}\int_2^{b+2} g(\pi(t))\,\mathrm{d}t = \frac{1}{2n\ell(n)}\int_{\mathrm{e}}^{n-1} t^2 \ell(t)\Delta(t)\,\mathrm{d}t + \frac{c_g}{n\ell(n)} \tag{2.4}$$

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

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

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

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

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

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 $n \geqslant 4$. Finally, we deduce (2.3) by applying the known approximations (see [2] and [1], respectively)

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

and

$$p_n = n(\ell(n) + O(1)), \qquad \text{as } n \to \infty, \tag{2.6}$$

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

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

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

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

Also, we have $p_n \leqslant b + 2 \leqslant p_{n+1}$, 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.      □

*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 Li is the logarithmic integral function defined by

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

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

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

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

$$\int \log\log x\,\mathrm{d}x = x\log\log x - \mathrm{Li}(x), \tag{3.2}$$

and this gives

$$\int \ell(x)\,\mathrm{d}x = \int \log(x\log x)\,\mathrm{d}x = x\log x + x\log\log x - x - \mathrm{Li}(x). \tag{3.3}$$

Moreover, by elementary computations, we have

$$\int \frac{\ell(x)}{x}\,\mathrm{d}x = \frac{1}{2}\log^2 x + \log x\log\log x - \log x. \tag{3.4}$$

*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

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

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

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

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

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

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, $\int_e^{n-1} t^2\Delta(t)\,\mathrm{d}t = O(\log n)$, and $\int_e^{n-1} t^2\ell(t)\Delta(t)\,\mathrm{d}t = \log^2 n + 2\log n\log\log n + O(\log n)$. Also, (with $g(x) = \frac{1}{x}$) we have $g(n) - \frac{n}{2}(g(n) - g(n-1)) = O(\frac{1}{n})$ and $g(n) + n(g(n) - g(n-1)) = O(\frac{1}{n})$. 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 $n \to \infty$, we obtain

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

and this gives (1.3). The proof is completed.                              □

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

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

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 $\eta \geqslant 0$, as well as $v(0) = 1$ and $\lim_{\eta\to\infty} v(\eta) = \infty$. Thus, for any real number $\beta \geqslant 1$ there exists a real number $\eta \geqslant 0$ such that $v(\eta) = \beta$, as desired.                □

## Acknowledgments

## 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.