On the Hardy-Ramanujan Theorem

Understanding global & local behaviour of prime omega function

Abhisruta Maity

3rd Year BS-MS

DMS Symposium · January 20, 2024

Department of Mathematics and Statistics

Indian Institute of Science Education and Research, Kolkata

Background

1. **Objects**: Arithmetic functions (sequences) $f : \mathbb{N} \to \mathbb{R}$. Ex: $f(n) = n^2, \log(n), d(n), \omega(n), \mu(n)$.

1. **Objects**: Arithmetic functions (sequences) $f : \mathbb{N} \to \mathbb{R}$. Ex: $f(n) = n^2, \log(n), d(n), \omega(n), \mu(n)$.
2. **Big *Oh-Notation***: For real-valued function f, a non-negative function g defined on $[a, \infty), a \geq 0$ and $x_0 \geq a$ we say

$$f(x) = O(g(x)) \text{ or } f(x) \ll g(x) \text{ for all } x \geq x_0,$$

if $\exists\, C > 0$ such that $\forall x \geq x_0$,

$$|f(x)| \leq Cg(x).$$

If x_0 is not specified, then the notation $f(x) \ll g(x)$ assumes the existence of some appropriately large positive real for which the above statement is true.

3. **Little *Oh*-Notation:** For real-valued function f, and a positive function g defined on $[a, \infty), a \geq 0$ we say

$$f(x) = o(g(x)) \text{ as } x \to \infty,$$

if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$

3. **Little *Oh*-Notation:** For real-valued function f, and a positive function g defined on $[a, \infty), a \geq 0$ we say

$$f(x) = o(g(x)) \text{ as } x \to \infty,$$

if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$

4. **Asymptotic $\sim$:** For positive real-valued function f and g defined on $[a, \infty), a \geq 0$ we say

$$f(x) \sim g(x) \text{ as } x \to \infty$$

if $f(x) = g(x) + o(g(x))$.

The Function $\omega(n)$

Define $\omega(n)$ to be *the number of distinct prime factors* of n

Define $\omega(n)$ to be *the number of distinct prime factors* of n:

$$\omega(n) := \sum_{p|n} 1,$$

for $n \in \mathbb{N}$.

Define $\omega(n)$ to be *the number of distinct prime factors* of n:

$$\omega(n) := \sum_{p \mid n} 1,$$

for $n \in \mathbb{N}$.

Example:

The Function $\omega(n)$

Define $\omega(n)$ to be *the number of distinct prime factors* of n:

$$\omega(n) := \sum_{p|n} 1,$$

for $n \in \mathbb{N}$.

Example:

- $\omega(1) = 0$

Define $\omega(n)$ to be *the number of distinct prime factors* of n:

$$\omega(n) := \sum_{p|n} 1,$$

for $n \in \mathbb{N}$.

Example:

- $\omega(1) = 0$
- $\omega(2) = 1$

Define $\omega(n)$ to be *the number of distinct prime factors* of n:

$$\omega(n) := \sum_{p \mid n} 1,$$

for $n \in \mathbb{N}$.

Example:

- $\omega(1) = 0$
- $\omega(2) = 1$
- $\omega(p^k) = 1$ for any prime p and $k \in \mathbb{N}$

The Function $\omega(n)$

Define $\omega(n)$ to be *the number of distinct prime factors* of n:

$$\omega(n) := \sum_{p|n} 1,$$

for $n \in \mathbb{N}$.

Example:

- $\omega(1) = 0$
- $\omega(2) = 1$
- $\omega(p^k) = 1$ for any prime p and $k \in \mathbb{N}$
- $\omega(200560490130) = \omega(2 \cdot 3 \cdot 5 \cdot \ldots \cdot 31) = 11$

The Function $\omega(n)$

Define $\omega(n)$ to be *the number of distinct prime factors* of n:

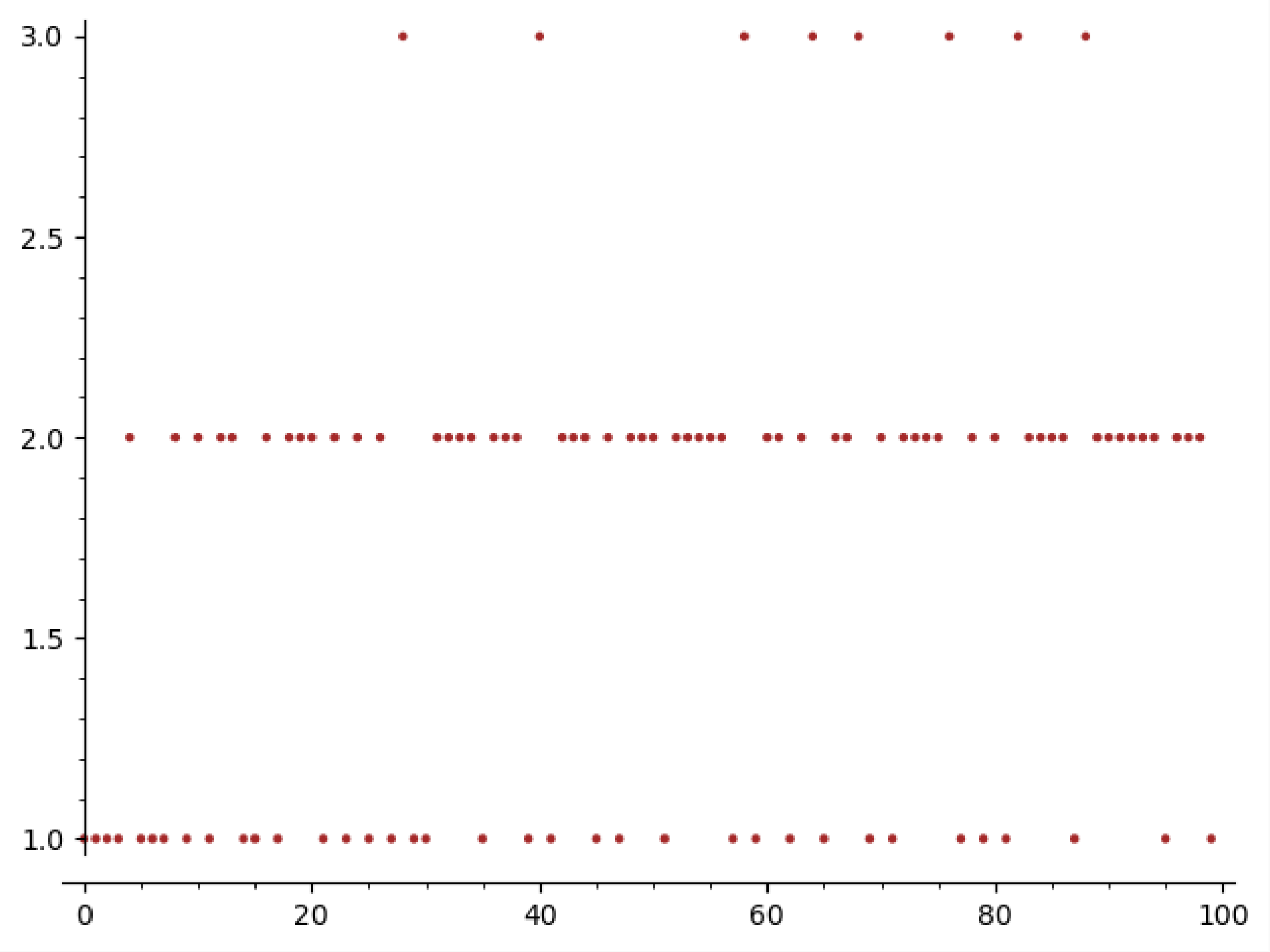
$$\omega(n) := \sum_{p|n} 1,$$

for $n \in \mathbb{N}$.

Example:

- $\omega(1) = 0$
- $\omega(2) = 1$
- $\omega(p^k) = 1$ for any prime p and $k \in \mathbb{N}$
- $\omega(200560490130) = \omega(2 \cdot 3 \cdot 5 \cdot ... \cdot 31) = 11$
- $\omega(200560490131) = 1$

Visual: Plot of $\omega(n)$



x : natural number n; y : $\omega(n)$

Aim

Understand the following for $\omega(n)$:

Aim

Understand the following for $\omega(n)$:

- Global behaviour in terms of *average order*.

Aim

Understand the following for $\omega(n)$:

- Global behaviour in terms of *average order*.
- Local behaviour via the notion of *normal order*.

Averages

A real-valued function g (generally monotonic) is called *average order* of the arithmetic function f, if

$$\frac{f(1)+f(2)+\ldots+f(N)}{N}=\frac{1}{N}\sum_{n\leq N}f(n)\sim g(N) \text{ as } N\to\infty.$$

Average of $\omega(n)$

Theorem 0.1: Average of $\omega(n)$ is $\log\log(n)$, $i.e.$,

$$E(N) = \frac{1}{N}\sum_{n\leq N}\omega(n) = \log\log N + O(1).$$

Theorem 0.2: Average of $\omega(n)$ is $\log\log(n)$, $i.e.$,

$$E(N) = \frac{1}{N}\sum_{n\leq N}\omega(n) = \log\log N + O(1).$$

Proof:

$$\sum_{n\leq N}\omega(n) = \sum_{n\leq N}\sum_{p|n}1 = \sum_{p\leq N}\sum_{\substack{n\leq N\\ n=pr}}1 = \sum_{p\leq N}\sum_{r\leq\frac{N}{p}}1 = \sum_{p\leq N}\left[\frac{N}{p}\right]$$

Average of $\omega(n)$

$$
\begin{aligned}
&= \sum_{p \le N} \left(\frac{N}{p} + O(1) \right) = N \sum_{p \le N} \frac{1}{p} + O(N) \\
&= N \sum_{p \le N} \frac{1}{p} + O(N) = N(\log\log N + O(1)) + O(N) \\
&= N \log\log N + O(N).
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{p \le N} \left(\frac{N}{p} + O(1) \right) = N \sum_{p \le N} \frac{1}{p} + O(N) \\
&= N \sum_{p \le N} \frac{1}{p} + O(N) = N(\log \log N + O(1)) + O(N) \\
&= N \log \log N + O(N).
\end{aligned}
$$

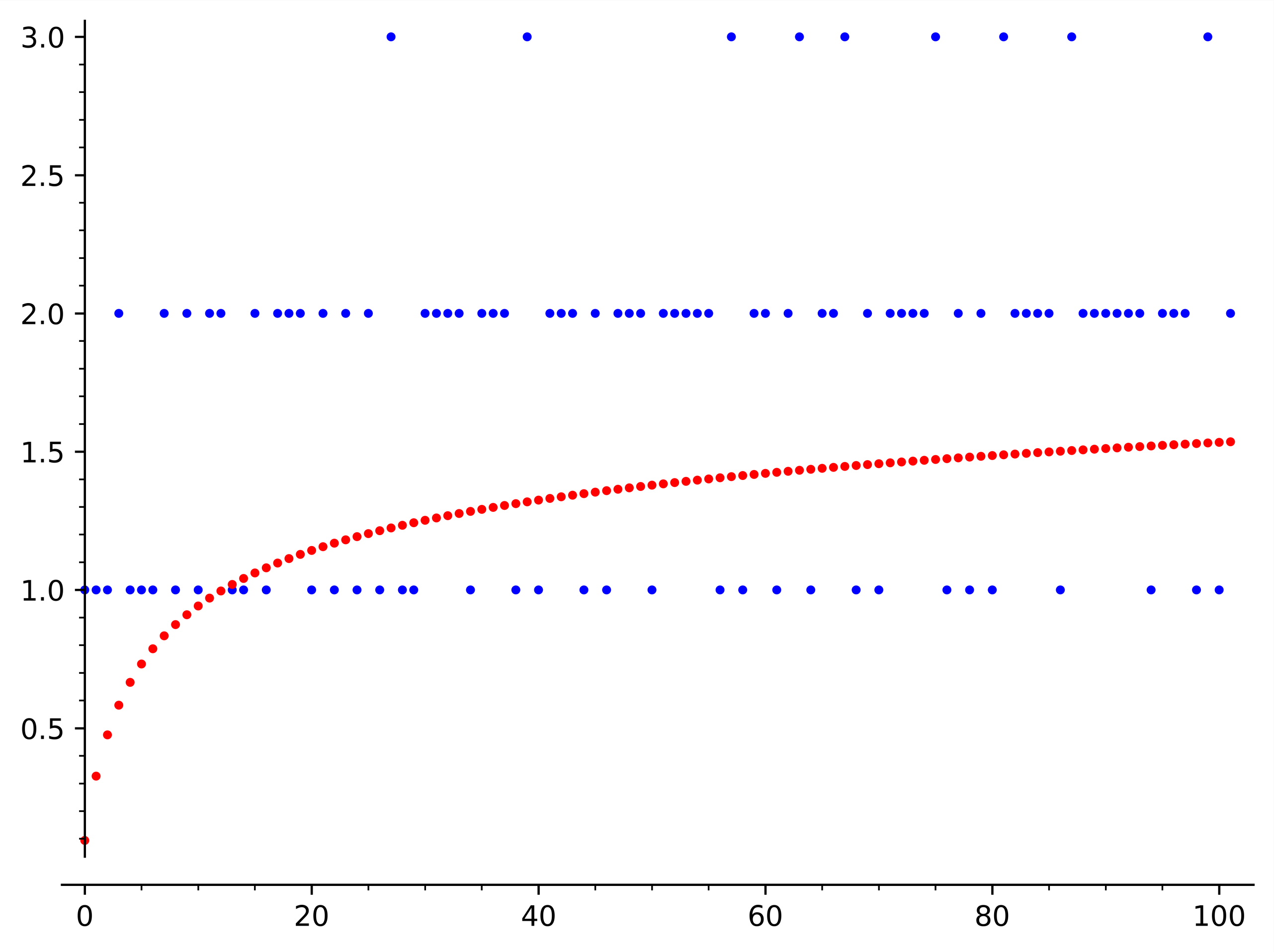
Therefore,

$$E(N) = \frac{1}{N} \sum_{n \le N} \omega(n) = \log \log N + O(1).$$

$\square$

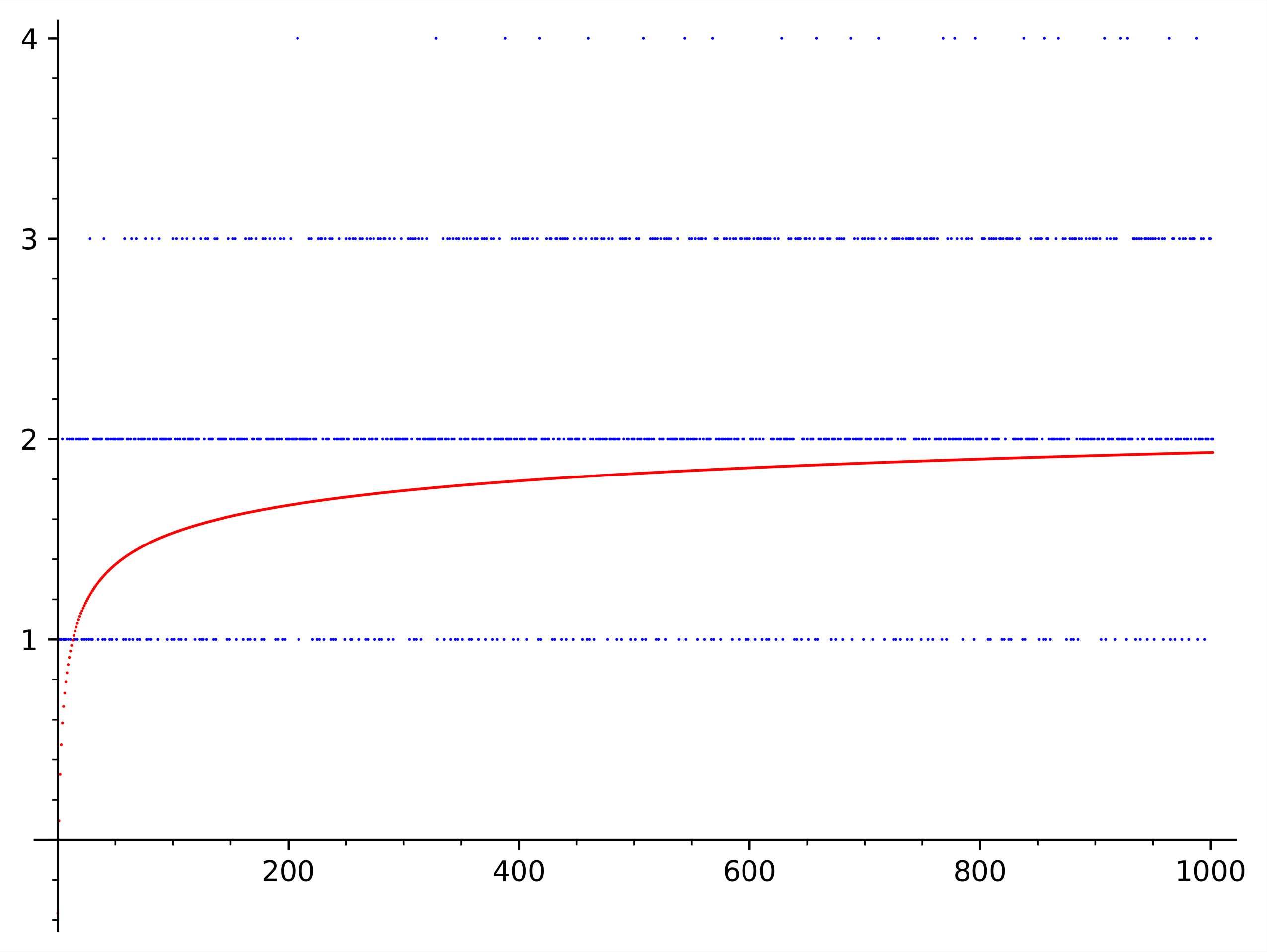
Deviation from Average

Visual: Deviation from Average



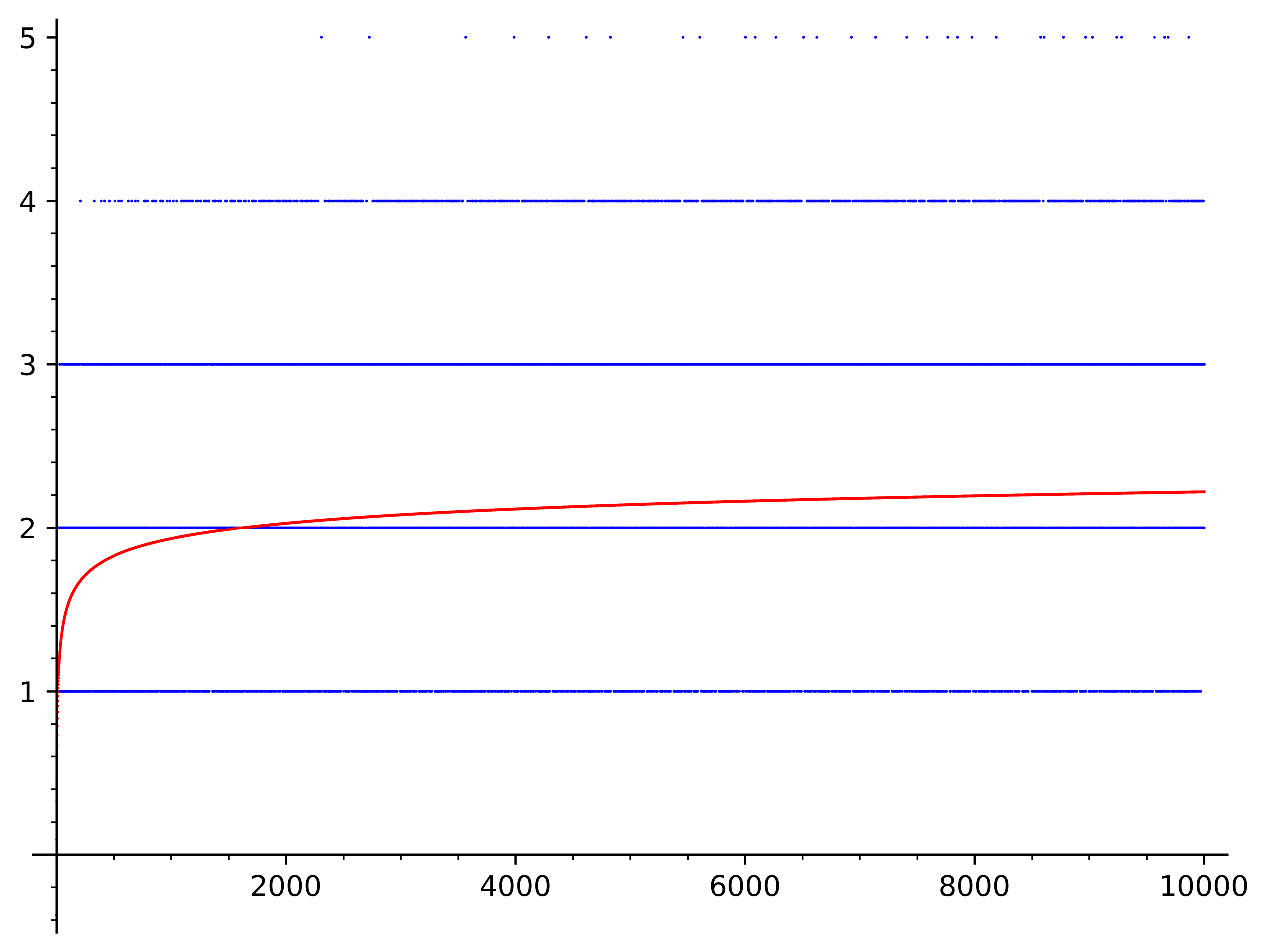
BLUE: $\omega(n)$; RED: $\log\log(n)$

Visual: Deviation from Average



BLUE: $\omega(n)$; RED: $\log\log(n)$

Visual: Deviation from Average



BLUE: $\omega(n)$; RED: $\log\log(n)$

One apparent candidate could be an estimation of *variance*:

$$V(N) = \frac{1}{N} \sum_{1<n\leq N} (\omega(n) - \log\log N)^2.$$

Variance

One apparent candidate could be an estimation of *variance*:

$$V(N) = \frac{1}{N} \sum_{1<n\leq N} (\omega(n) - \log\log N)^2.$$

In 1943, *Pál Turán* showed:

$$V(N) = O(\log\log N).$$

Significance of Turán's Estimate

- Shows variance is of same order as the average.

Significance of Turán's Estimate

- Shows variance is of same order as the average.
- Motivates to calculate higher moments and ...

Proof of Turán's Estimate

$$\frac{1}{N}\sum_{1<n\leq N}(\omega(n)-\log\log(N))^2$$

$$=\frac{1}{N}\underbrace{\sum_{1<n\leq N}\omega(n)^2}_{S(N)}-2\log\log(N)\underbrace{\frac{1}{N}\sum_{1<n\leq N}\omega(n)}_{\log\log(N)+O(1)}+(\log\log(N))^2\frac{1}{N}\underbrace{\sum_{n\leq N}1}_{N+O(1)}$$

$$=\frac{1}{N}\sum_{1<n\leq N}\omega(n)^2$$

$$-2(\log\log(N))^2+O(\log\log(N))$$

$$+(\log\log(N))^2+O\left(\frac{(\log\log(N))^2}{N}\right)$$

$$=\frac{1}{N}\sum_{1<n\leq N}\omega(n)^2-(\log\log N)^2+O(\log\log N)$$

$$S(N) = \sum_{1<n\leq N} \omega(n)^2 = \sum_{1<n\leq N} \sum_{p|n} 1 \sum_{q|n} 1 = \sum_{1<n\leq N} \sum_{p|n} \sum_{q|n} 1,$$

which can be broken into following two sums: $S_1(N) = \sum_{1<n\leq N} \sum_{p|n} 1$ and $S_2(N) = \sum_{1<n\leq N} \sum_{\substack{pq|n \\ p\neq q}} 1$.

$$S_1(N) = \sum_{1<n\leq N} \sum_{p|n} 1 = \sum_{1<n\leq N} \omega(n)$$
$$= N \log\log(N) + O(N).$$

Proof of Turán's Estimate

$$
\begin{aligned}
S_2(N) &= \sum_{\substack{pq \le N \\ p \ne q}} \sum_{r : n = rpq \le N} 1 \\
&= \sum_{\substack{pq \le N \\ p \ne q}} \left[\frac{N}{pq} \right] \\
&= N \sum_{\substack{pq \le N \\ p \ne q}} \frac{1}{pq} + O\left(\sum_{\substack{pq \le N \\ p \ne q}} 1 \right).
\end{aligned}
$$

Proof of Turán's Estimate

$$S_3(N) = \sum_{\substack{pq \leq N \\ p \neq q}} 1$$

$$= \sum_{p \leq \sqrt{N}} \sum_{q \leq \frac{N}{p}} 1 \ll \sum_{p \leq \sqrt{N}} \left[\frac{N}{p} \right] \leq \sum_{p \leq \sqrt{N}} \frac{N}{p} \ll N \log\log(N),$$

Then

$$S_2(N) = N \sum_{\substack{pq \leq N \\ p \neq q}} \frac{1}{pq} + O(N \log\log(N)),$$

Proof of Turán's Estimate

$$S_4(x) = N \sum_{\substack{pq \le N \\ p \ne q}} \frac{1}{pq} = N \sum_{pq \le N} \frac{1}{pq} - N \underbrace{\sum_{p^2 \le N} \frac{1}{p^2}}_{O(1)} = N \sum_{pq \le N} \frac{1}{pq} + O(1).$$

Observe that,

$$\left(\sum_{p \le \sqrt{N}} \frac{1}{p} \right)^2 \le \sum_{pq \le N} \frac{1}{pq} \le \left(\sum_{p \le N} \frac{1}{p} \right)^2,$$

which implies

$$\left(\log\log\left(\sqrt{N}\right) + R_1(N) \right)^2 \le \sum_{pq \le N} \frac{1}{pq} \le \left(\log\log(N) + R_2(N)\right)^2,$$

where $R_1(N), R_2(N)$ both are $O(1)$. The above implies

$$\sum_{pq \leq N} \frac{1}{pq} = (\log\log(N))^2 + O(\log\log(N)).$$

Therefore,

$$S_2(N) = N(\log\log(N))^2 + O(N\log\log(N)),$$

which plugging back yields,

$$S(N) = N(\log\log(N))^2 + O(N\log\log(N)).$$

A Small Corollary

If we estimate the total pointwise deviation:

$$\tilde{V}(N) = \frac{1}{N} \sum_{1<n\leq N} (\omega(n) - \log\log(n))^2,$$

If we estimate the total pointwise deviation:

$$\tilde{V}(N) = \frac{1}{N} \sum_{1<n\leq N} (\omega(n) - \log\log(n))^2,$$

then again the estimate doesn't differ too much:

$$\tilde{V}(N) = O(\log\log N).$$

A Small Corollary

If we estimate the total pointwise deviation:

$$\tilde{V}(N) = \frac{1}{N} \sum_{1<n\leq N} (\omega(n) - \log\log(n))^2,$$

then again the estimate doesn't differ too much:

$$\tilde{V}(N) = O(\log\log N).$$

Because:

- Upto $N^{\frac{1}{e}}$ there are $o(N)$ number of numbers & $\omega(n) \ll \log N$.

A Small Corollary

If we estimate the total pointwise deviation:

$$\tilde{V}(N) = \frac{1}{N} \sum_{1<n\leq N} (\omega(n) - \log\log(n))^2,$$

then again the estimate doesn't differ too much:

$$\tilde{V}(N) = O(\log\log N).$$

Because:

- Upto $N^{\frac{1}{e}}$ there are $o(N)$ number of numbers & $\omega(n) \ll \log N$.
- For $N^{\frac{1}{e}} \leq n \leq N$,

$$\log\log(N) - 1 \leq \log\log(n) \leq \log\log(N)$$

Normal Order

A positive real-valued function g is called *normal order* of the arithmetic function f, if for every $\varepsilon > 0$,

$$\#\left\{n \leq N : \left|\frac{f(n)}{g(n)} - 1\right| \geq \varepsilon\right\} = o(N) \text{ as } N \to \infty.$$

We also say

$$f(n) \sim g(n) \qquad \text{a.e.}$$

to convey the same notion.

Theorem 0.3: For every $0 < \delta < \frac{1}{2}$,

$$\#\underbrace{\left\{1 < n \leq N : \left|\frac{\omega(n)}{\log\log(n)} - 1\right| \geq \frac{1}{(\log\log(n))^{\frac{1}{2}-\delta}}\right\}}_{=:B_\delta(N)} = o(N).$$

Hardy-Ramanujan Theorem (1917)

Theorem 0.4: For every $0 < \delta < \frac{1}{2}$,

$$\#\underbrace{\left\{1 < n \leq N : \left|\frac{\omega(n)}{\log\log(n)} - 1\right| \geq \frac{1}{(\log\log(n))^{\frac{1}{2}-\delta}}\right\}}_{=: B_\delta(N)} = o(N).$$

- RHS becomes small for sufficiently large n.

Hardy-Ramanujan Theorem (1917)

Theorem 0.5: For every $0 < \delta < \frac{1}{2}$,

$$\#\underbrace{\left\{1 < n \leq N : \left|\frac{\omega(n)}{\log\log(n)} - 1\right| \geq \frac{1}{(\log\log(n))^{\frac{1}{2}-\delta}}\right\}}_{=: B_\delta(N)} = o(N).$$

- RHS becomes small for sufficiently large n.

Corollary 0.5.1: The normal order of $\omega(n)$ is $\log\log(n)$.

Why it's important?

- Looks like Chebyshev's inequality in probability. Opens up the question: whether results of probability can be used to solve number theoretic problems.

Why it's important?

- Looks like Chebyshev's inequality in probability. Opens up the question: whether results of probability can be used to solve number theoretic problems.
- Whether this can be generalized for other additive functions.

Why it's important?

- Looks like Chebyshev's inequality in probability. Opens up the question: whether results of probability can be used to solve number theoretic problems.
- Whether this can be generalized for other additive functions.
- Computationally, the approximation is good enough in many situations.

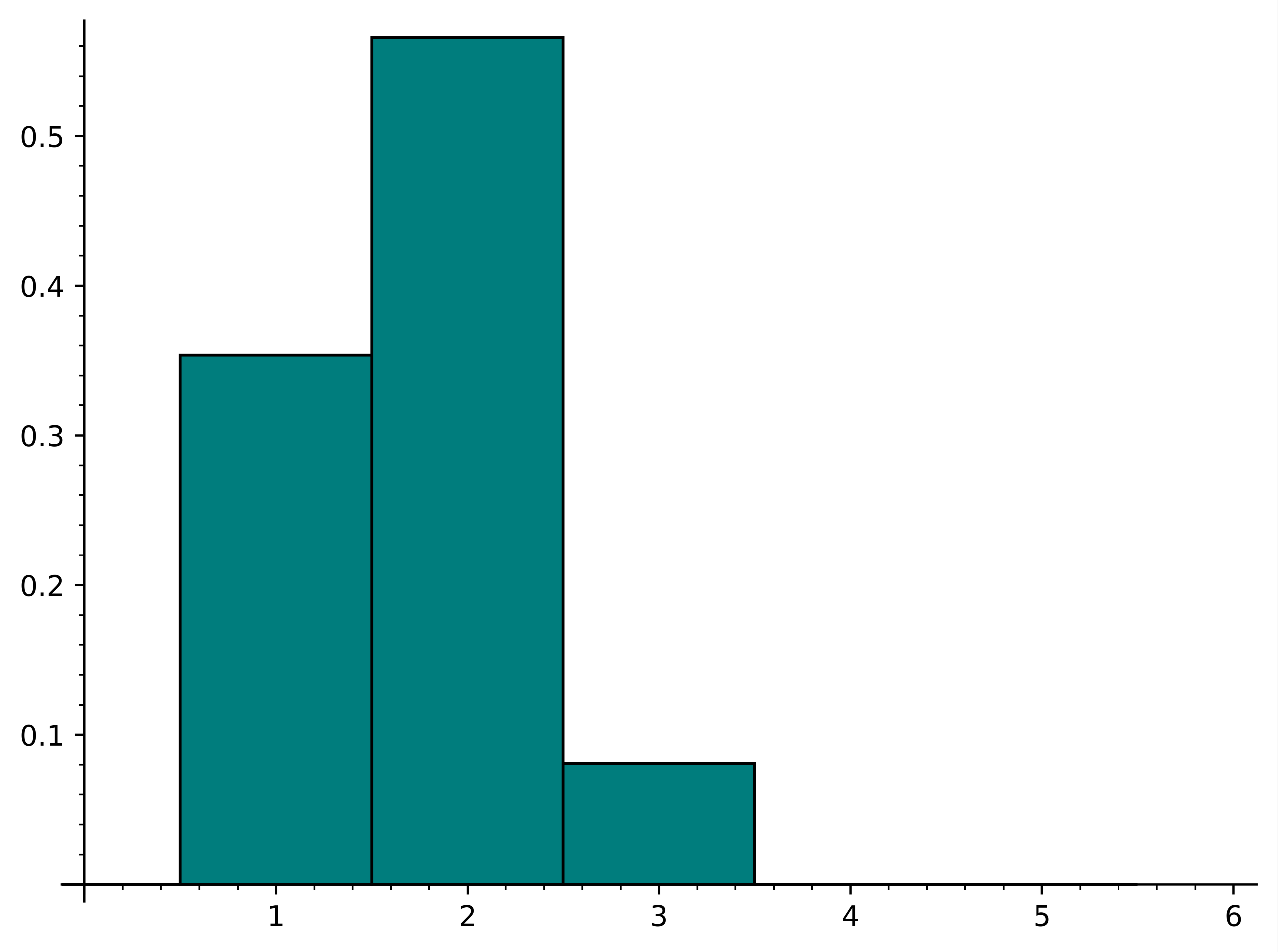
Proof

Suppose to the contrary that for some $\delta > 0$ there is an $\alpha > 0$, such that for each $N \in \mathbb{N}$, $\#B_\delta(N) \geq \alpha N$. For any large N, then:

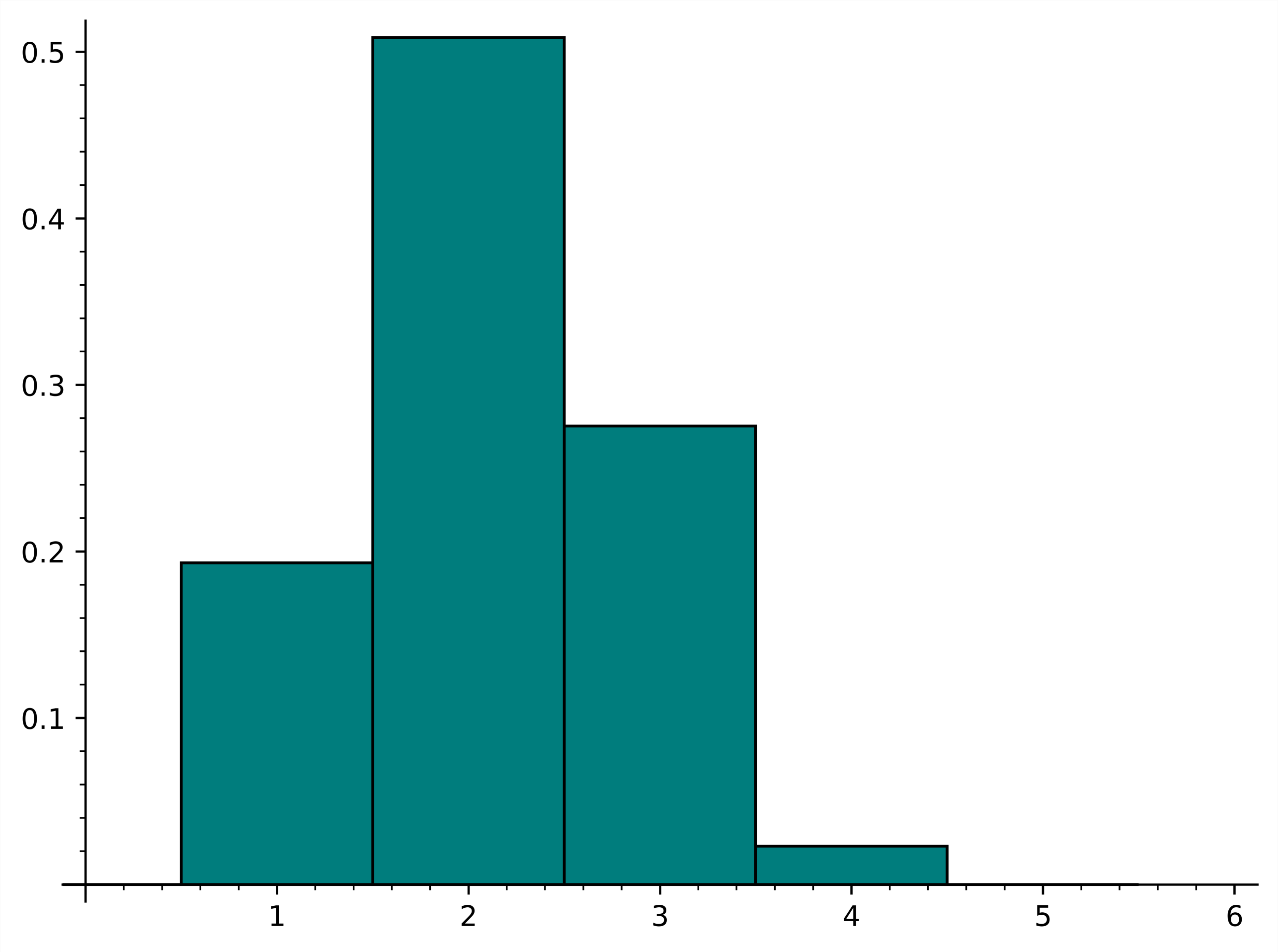
$$\begin{aligned}
\tilde{V}(N) &= \frac{1}{N} \sum_{1<n\leq N} \left(\omega(n) - \log\log(n)\right)^2 \\
&\geq \frac{1}{N} \sum_{N^{\frac{1}{e}}<n\leq N} \left(\omega(n) - \log\log(n)\right)^2 \\
&\geq \frac{1}{N} \alpha N \left(\log\log(N) - 1\right)^{1+2\delta} \\
&= \alpha' \left(\log\log(N)\right)^{1+2\delta} \text{ for some } \alpha' > 0,
\end{aligned}$$

which is a contradiction since $\tilde{V}(N) = O(\log\log N)$.

Further ambition

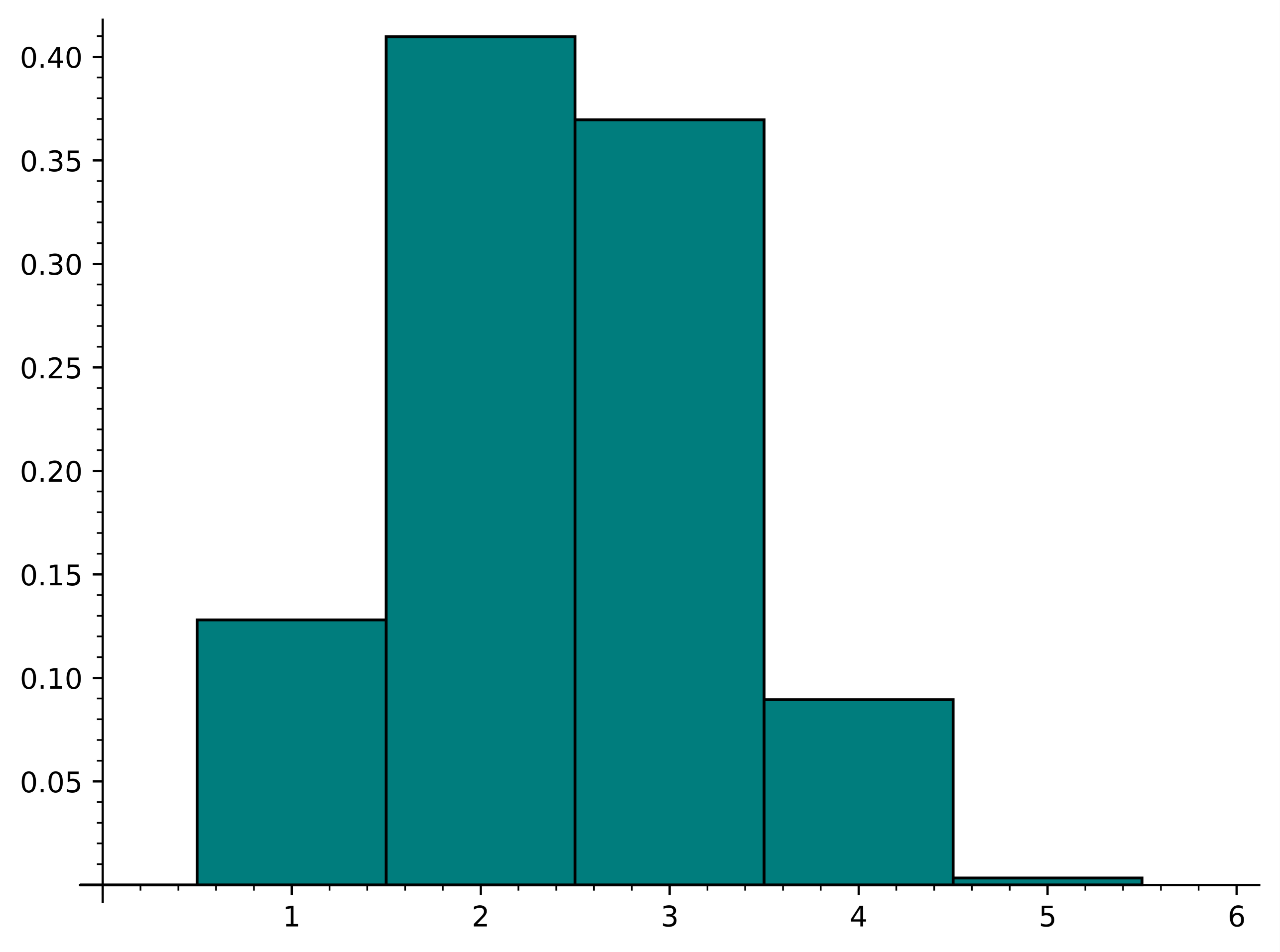


x : ω value; y : density with $N = 10^2$



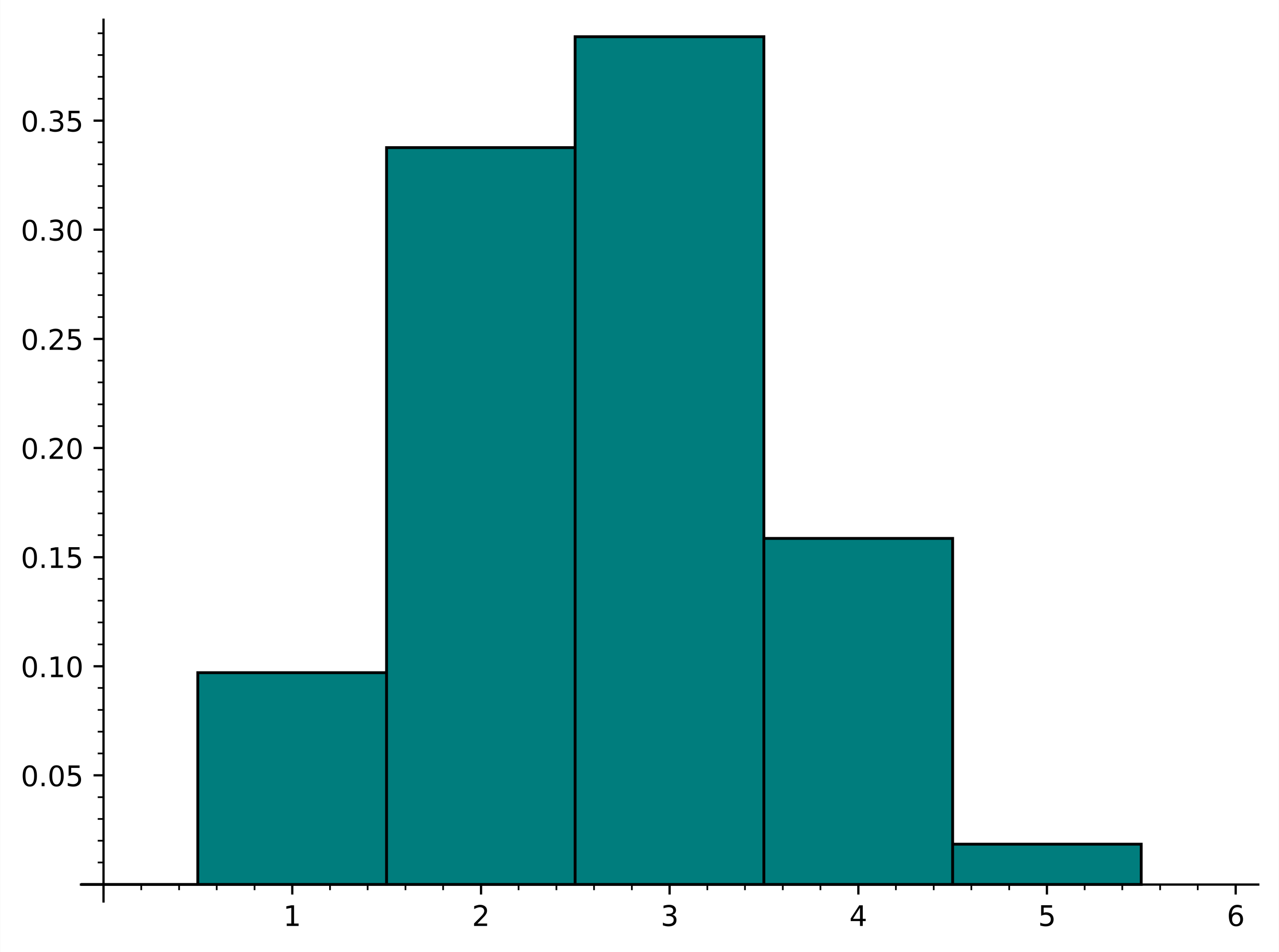
x : ω value; y : density with $N = 10^3$

Further ambition



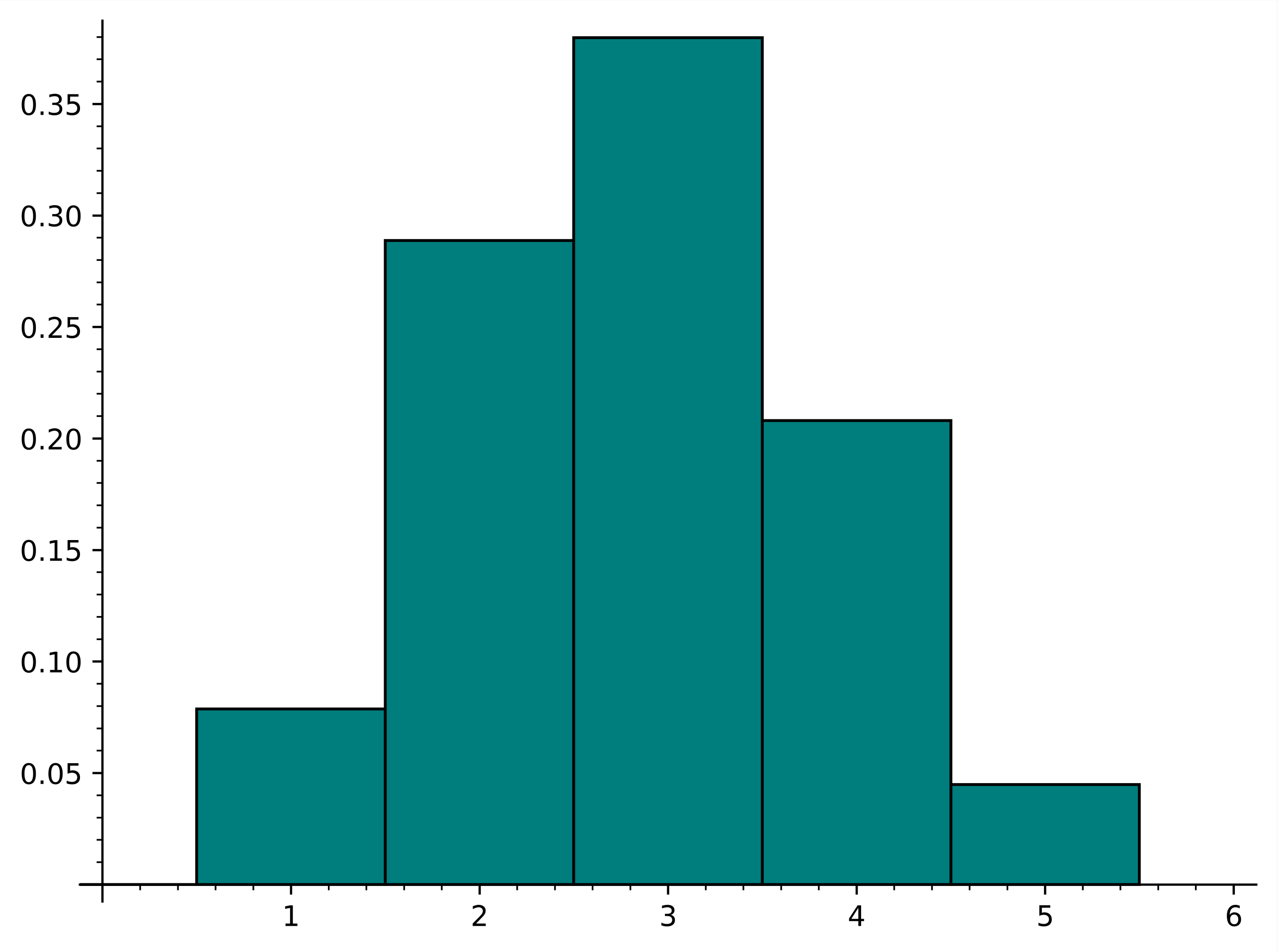
x : ω value; y : density with $N = 10^4$

Further ambition



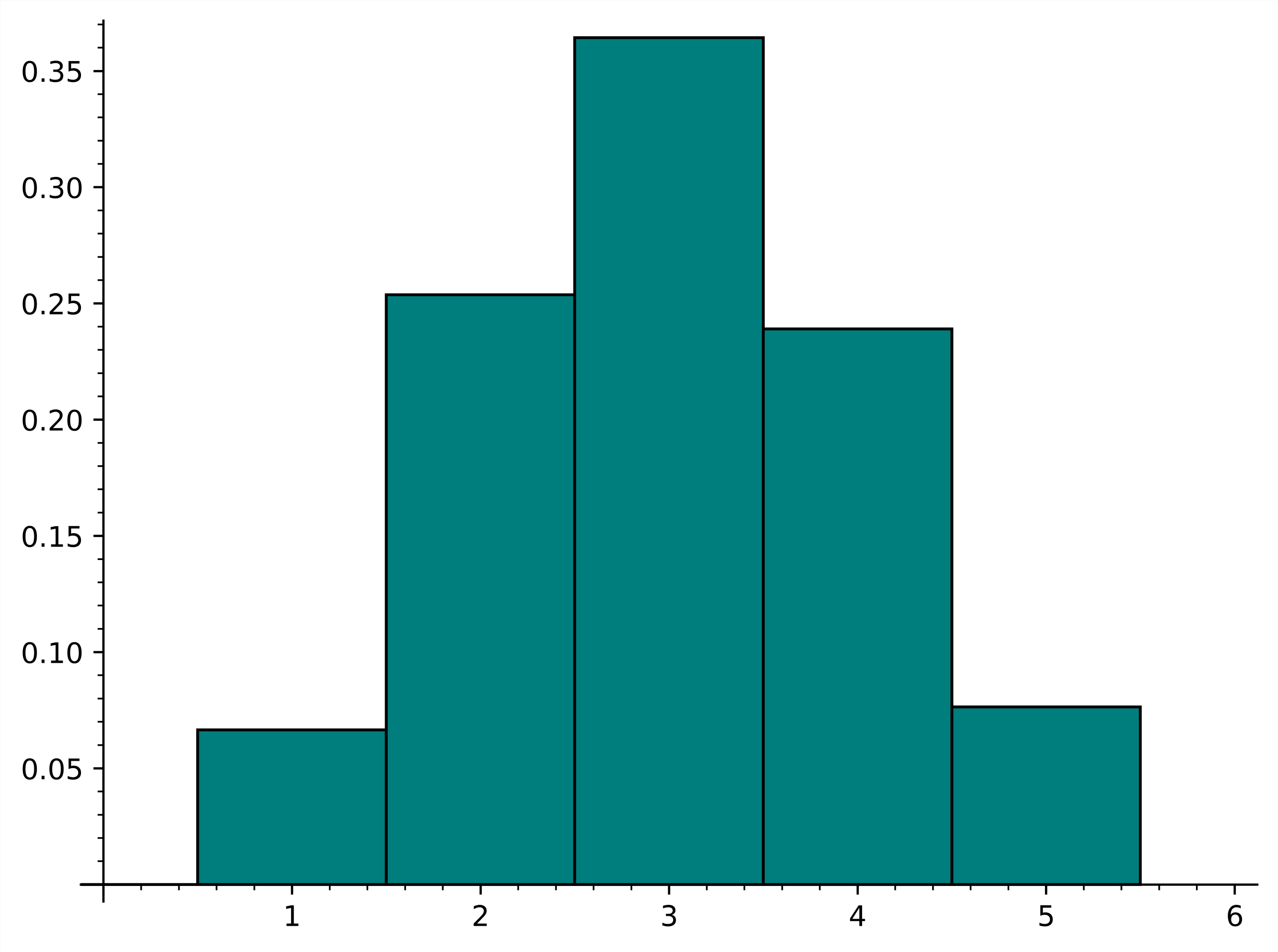
x : ω value; y : density with $N = 10^5$

Further ambition



x : ω value; y : density with $N = 10^6$

Further ambition



x : ω value; y : density with $N = 10^7$

‘Normal’ order makes sense!

The elegant Erdős-Kac Theorem:

$$\lim_{N\to\infty} \frac{1}{N}\#\left\{n \le N : \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \le x\right\} = \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt,$$

which is called the Central Limit Theorem of probabilistic number theory.

Thank you!

References:

1. De Koninck, J.M. and Luca, F. (2023). Analytic number theory: Exploring the anatomy of integers (Vol. 134). American Mathematical Society.
2. Cojocaru, A. C. and Murty, M. R. (2005) An Introduction to Sieve Methods and Their Applications. Cambridge: Cambridge University Press (London Mathematical Society Student Texts).
3. Ramanujan, S. and Hardy, G. (1917). The normal number of prime factors of a number n. Quarterly Jour, of Math, 48, pp.76-92.
4. Turán, P. (1934). On a theorem of Hardy and Ramanujan. Journal of the London mathematical society, 1(4), pp.274-276.
5. Lectures in Probabilistic Number Theory by Kaneenika Sinha.