#A80 INTEGERS 20 (2020)

AN ERDŐS-KAC THEOREM FOR SMOOTH AND ULTRA-SMOOTH INTEGERS

Marzieh Mehdizadeh
Dépt. de Mathématiques et Statistique, Université de Montréal, Québec, Canada
marzieh.mehdizadeh@gmail.com

Received: 10/5/17, Revised: 4/12/20, Accepted: 9/8/20, Published: 10/9/20

Abstract

The Erdős-Kac theorem is an extension of the Hardy-Ramanujan theorem which states that, if $\omega(n)$ is the number of distinct prime factors of an integer n, the normal order of $\omega(n)$ is $\log\log n$ with a typical error of size $\sqrt{\log\log n}$. Here we prove an Erdős-Kac type of theorem for the set $S(x,y) := \{n \le x : p|n \Rightarrow p \le y\}$. In fact, we prove that the distribution of $\omega(n)$ for $n \in S(x,y)$ is Gaussian in a certain range of y by using the method of moments. The advantage of this approach is to recover classical results in the range $u = o(\log\log y)$, where $u := \frac{\log x}{\log y}$, by introducing a much simpler proof.

1. Introduction

For an integer $n \ge 2$, let $\omega(n)$ be the number of distinct prime divisors of n. In 1940, Erdős and Kac [5] in their celebrated work, studied the distribution of $\omega(n)$ in the interval $[2, N]$, for a large enough positive integer N. The theorem states that for any real number x, we have

$$\lim_{N\to\infty} \frac{1}{N} \# \left\{ n \le N : \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \le x \right\} = \Phi(x), \tag{1}$$

where $\Phi(x)$ is the cumulative distribution function (CDF) of the standard normal distribution denoted by

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

There are several proofs of the Erdős-Kac theorem. For instance, the Erdős-Kac theorem has been proved by Billingsley [2] using the method of moments, and by Granville and Soundararajan [7] using sieve theory. Also, different variations of this theorem have been considered by several authors.

In the present note, we shall study the Erdős-Kac theorem for y-smooth integers. Recall that

$$S(x, y) := \{n \leq x : P(n) \leq y\}, \quad x \geq y \geq 2,$$

is the set of y-smooth integers less than x, where $P(n)$ is defined as the largest prime factor of n, with the convention $P(1) = 1$. Also, recall that we set

$$\Psi(x, y) := |S(x, y)|, \quad x \geq y \geq 2.$$

The main goal is to prove a similar result to (1) where we only consider the integers in $S(x, y)$, where x and y satisfy the range

$$u = o(\log \log y), \tag{2}$$

and u is defined by

$$u := \frac{\log x}{\log y}.$$

It is good to give a history about the Erdős-Kac theorem and the different approaches to solve it. Hildebrand [9], Alladi [1], and Hensley [8] have considered the distribution of prime divisors of y-smooth integers in different ranges of y. Hensley proved an Erdős-Kac type theorem when u lies in the range

$$(\log y)^{1/3} \leq u \leq \frac{\sqrt{y}}{2 \log y}.$$

By using a different method, Alladi obtained another type of the Erdős-Kac theorem in the range

$$u \leq \exp(\log y)^{3/5-\epsilon}.$$

Later, Hildebrand extended previous results to include the range

$$y \geq 3, \qquad u \geq (\log y)^{20},$$

where the above range is a completion of the ranges in Alladi and Hensley's results. Although (2) does not cover Alladi's, Hensley's and Hildebrand's ranges, our new approach is completely different and much simpler than the methods used by previous authors. Our approach is based on the method of moments in [2] to solve these sort of problems by an easier probabilistic method. The methodology is to define a family of independent normally distributed random variables with the same variance and mean value as $\omega(n)$ (where $n \in S(x, y)$), and by applying the method of moments one can get the desired result in (1). The first step of our proof is to remove the big prime divisors which are negligible on the proof process. Therefore, we apply a truncation on the number of prime factors of a typical integer n in $S(x, y)$. The idea comes from the original proof of the Erdős-Kac theorem [5]. For a given real number y, set

$$\phi(y) := (\log \log y)^{\sqrt{\log \log \log y}}.$$

We will see that $y^{\frac{1}{\phi(y)}}$ is a function that helps us to sieve out all primes exceeding $y^{\frac{1}{\phi(y)}}$, and we shall show that the contribution of sieved primes is negligible in understanding the distribution of $\omega(n)$. Before stating the main result, we begin introducing some notation. If $\omega(n)$ is the number of distinct prime divisors of a y-smooth integer, namely

$$\omega(n) := \sum \mathbb{1}_{p|n}(n),$$

where $n \in S(x,y)$ and $\mathbb{1}_{p|n}(n)$ is 1 or 0 according to whether or not the prime p divides n. Let $\mu_\omega(x,y)$ be the mean value of $\omega(n)$, more formally

$$\mu_\omega(x,y) := \mathbb{E}_{n\in S(x,y)}[\omega(n)] = \frac{1}{\Psi(x,y)} \sum_{n\in S(x,y)} \omega(n),$$

and let $\sigma_\omega^2(x,y)$ be the variance of $\omega(n)$, defined by

$$\sigma_\omega^2(x,y) := \mathbb{E}\left[(\omega(n) - \mu(x,y))^2\right].$$

Now we are ready to formulate the main theorem.

Theorem 1.1. *For any real value z, we have*

$$\frac{1}{\Psi(x,y)}\#\{n \in S(x,y) : \frac{\omega(n) - \log\log y}{\sqrt{\log\log y}} \le z\} \to \Phi(z) \quad as \quad y \to \infty, \tag{3}$$

holds in range (2).

Theorem 1.1 is proved in Section 3. The proof relies on estimating $\Psi(x/d,y)/\Psi(x,y)$, where $d \le y$, and the method of moments. Let $U(x,y)$ be the set of y-ultra-smooth integers whose canonical decomposition is free of prime powers exceeding y, more formally

$$U(x,y) := \{n \le x : p^v || n \Rightarrow v \le v_p\},$$

where

$$v_p := \left\lfloor \frac{\log y}{\log p} \right\rfloor.$$

We denote the number of y-ultra-smooth integers as follows:

$$\Upsilon(x,y) := \left|U(x,y)\right|,$$

and we now are ready to state our second theorem.

Theorem 1.2. *For any real number z, we have*

$$\frac{1}{\Upsilon(x,y)}\#\{n \in U(x,y) : \frac{\omega(n) - \log\log y}{\sqrt{\log\log y}} \le z\} \to \Phi(z) \quad as \quad y \to \infty, \tag{4}$$

holds in the range (2).

The proof of Theorem 1.2 relies on the method of moments and local behaviour of the function $\Upsilon(x,y)$. By recalling [11, Corollary 1.3.], for $u = o(\log\log y)$, we have

$$\frac{\Upsilon(x/d,y)}{\Upsilon(x,y)} = \frac{\Psi(x/d,y)}{\Psi(x,y)}\left\{1 + O\left(\frac{u\log 2u}{\sqrt{y}\log y}\right)\right\},$$

that is,

$$\frac{\Upsilon(x/d,y)}{\Upsilon(x,y)} \sim \frac{\Psi(x/d,y)}{\Psi(x,y)} \quad \text{as} \quad y \to \infty.$$

This relation between the local behaviour of $\Upsilon(x,y)$ and $\Psi(x,y)$ leads the reader to obtain a similar proof as the proof of Theorem 1.1, so we avoid proving Theorem 1.2.

2. Preliminaries

Here we briefly state a few useful standard facts in probability theory that are presented in [2] (See Feller [6] for more information), and we shall provide some important lemmas.

Remark 1. If a random variable D_n converges to 0 in probability, particularly $\mathbb{E}\{|D_n|\} \to 0$, then a second random variable U_n (on the same probability space) tends to Φ in distribution if and only if $U_n + D_n \to \Phi$ in distribution.

Remark 2. If distribution function F_n satisfies $\int_{-\infty}^{\infty} x^k dF_n(x) \to \int_{-\infty}^{\infty} x^k d\Phi(x)$ as $n \to \infty$, for $k = 1, 2, \cdots$, then $F_n(x) \to \Phi(x)$ for each x.

Remark 3. If $F_n(x) \to \Phi(x)$ for each x (point-wise), and if $\int_{-\infty}^{\infty} |x|^{k+\epsilon} dF_n(x)$ is bounded in n for some positive ϵ, then

$$\int_{-\infty}^{\infty} x^k dF_n(x) \to \int_{-\infty}^{\infty} x^k d\Phi(x).$$

The following important remark is a particular type of Central Limit Theorem(CLT).

Remark 4. If $X_1, X_2, \cdots$ are independent and uniformly bounded random variables with mean 0 and finite variance σ_i^2, and if $\sum \sigma_i^2$ diverges, then the distribution of $\frac{\sum_{i=1}^n X_i}{\left(\sum_{i=1}^n \sigma_i^2\right)^{1/2}}$ converges to the normal distribution function.

We now state some famous estimations about the local behaviour of $\Psi(x,y)$. By recalling [4, Theorem 2.4.], for $m = 1$, $d \le y$ and $y \ge (\log x)^{1+\epsilon}$, we have

$$\Psi(x/d,y) = \frac{\Psi(x,y)}{d^{\alpha}}\left\{1 + O\left(\frac{1}{u_y} + \frac{\log d}{\log x}\right)\right\}, \tag{5}$$

where $u_y := u + \frac{\log y}{\log(u+2)}$ and $\alpha = \alpha(x, y)$ denotes the saddle point of Perron Integral for $\Psi(x, y)$, which is the solution of the equation

$$\sum_{p \le y} \frac{\log p}{p^\alpha - 1} = \log x.$$

Since $\alpha(x, y)$ plays an important role in proving our theorems, we mention some fundamental facts about this function. By [4, Lemma 3.1], for any $\epsilon > 0$, we have

$$\alpha = 1 - \frac{\xi(u)}{\log y} + O\left(\frac{1}{L_\epsilon(y)} + \frac{1}{u(\log y)^2}\right), \quad \text{for} \quad y \ge (\log x)^{1+\epsilon}, \tag{6}$$

where $\xi(u)$ is a unique real non-zero root of the equation

$$e^\xi = 1 + u\xi,$$

and for $u \ge 3$, we have

$$\xi(u) = \log(u \log u) + O\left(\frac{\log \log u}{\log u}\right). \tag{7}$$

The following lemma is stated in [4].

Lemma 2.1. *[De la Brèteche, Tenenbaum]. For any $x \ge y \ge 2$, uniformly we have*

$$\sum_{p \le y} \frac{1}{p^\alpha} = \log \log y + \left\{1 + O\left(\frac{1}{\log y}\right)\right\} \frac{uy}{y + \log x}. \tag{8}$$

Here, we use a particular case of Lemma 2.1 when the range is restricted to $\log x < y \le x$. Then

$$\frac{uy}{y + \log x} = u\left(1 + O\left(\frac{\log x}{y}\right)\right),$$

and we obtain

$$\sum_{p \le y} \frac{1}{p^\alpha} = \log \log y + u + O\left(\frac{u}{\log y}\right), \qquad \text{for} \quad y > \log x. \tag{9}$$

We now define the truncated version $\omega(n)$ that helps us to prove our main theorem. Let t be in $2 \le t \le y \le x$, then we define

$$\omega_t(n) := \#\{p : p|n, p \le t\} = \sum_{p \le t} \mathbb{1}_{p|n}.$$

By applying the saddle point method, Tenenbaum and de la Brèteche [3] obtained an estimate for the expectation and the variance of $\omega_t(n)$. Before presenting that, we define

$$M(t) = M_{x,y}(t) := \sum_{p \le t} \frac{1}{p^\alpha}.$$

Now we are ready to state Lemma 9.1 from [3].

Lemma 2.2. *[Tenenbaum, de la Brèteche]. For $2 \le t \le y \le x$ we have uniformly*

$$\mu_{\omega_t}(x, y) = M(t) + O(1). \tag{10}$$

Now we estimate the expectation value of $\omega(n)$ in the range $u = o(\log\log y)$, where $n \in S(x, y)$.

Lemma 2.3. *If $u = o(\log\log y)$, then we have*

$$\mu_\omega(x, y) = \log\log y + o(\log\log y).$$

Proof. Set $t = y$ in Lemma 2.2. Then we have

$$\mu_\omega(x, y) = \sum_{p \le y} \frac{1}{p^\alpha} + O(1).$$

By using (9), we get

$$\mu_\omega(x, y) = \log\log y + u + O(1).$$

By setting $u = o(\log\log y)$, one can obtain

$$\mu_\omega(x, y) = \log\log y + o(\log\log y),$$

and the proof is complete. □

Lemma 2.4. *If $u = o(\log\log y)$ and $t \le y^{1/\log u}$, then we have*

$$\sum_{p \le t} \frac{1}{p^\alpha} = \log\log t + O(1). \tag{11}$$

Proof. Since $(1-\alpha)\log p$ is bounded in our range, we can rewrite the above summation as follows:

$$\sum_{p \le t} \frac{1}{p^\alpha} = \sum_{p \le t} \frac{1}{pp^{\alpha-1}} = \sum_{p \le t} \frac{1}{p} \left\{1 + O\left((1-\alpha)\log p\right)\right\}.$$

By using the given estimate for α in (6) and Mertens' estimate, one can get

$$\begin{aligned} \sum_{p \le t} \frac{1}{p^\alpha} &= \sum_{p \le t} \frac{1}{p} + O\left(\frac{\xi(u)}{\log y} \sum_{p \le t} \frac{\log p}{p}\right) \\ &= \log\log t + O\left(\frac{\xi(u)}{\log y} \log t\right). \end{aligned} \tag{12}$$

Finally, by applying the estimation for $\xi(u)$ in (7), we get our desired result. □

In the next lemma and corollary we show that the contribution of large prime factors of $n \in S(x,y)$ does not affect the expected value of the number of prime factors of n and hence the distribution of $\omega(n)$, where u is small enough. We define

$$\omega_Y(n) := \sum_{p\le Y} \mathbb{1}_{p|n}(y), \tag{13}$$

where

$$Y := y^{\frac{1}{\phi(y)}}, \quad \text{and} \quad \phi(y) := (\log\log y)^{\sqrt{\log\log\log y}}.$$

The above truncation helps us to claim the following lemma.

Lemma 2.5. *If $u = o(\log\log y)$, then we have*

$$\sum_{p\le Y} \frac{1}{p^\alpha} = \log\log y + O\left((\log\log\log y)^{3/2}\right).$$

Proof. The proof is easy and stems from Lemma 2.4:

$$\begin{aligned}\sum_{p\le Y} \frac{1}{p^\alpha} &= \log\log y - \log\phi(y) + O(1)\\ &= \log\log y - (\log\log\log y)^{3/2} + O(1),\end{aligned} \tag{14}$$

and we have our desired result. □

We now define the average of $\omega_Y(n)$ as follows:

$$\mu_{\omega_Y}(x,y) := \mathbb{E}\left[\omega_Y(n)\right].$$

By the following lemma, we will show that $\omega(n)$ can be replaced by its truncated version $\omega_Y(n)$ in the statement of Theorem 1.1.

Lemma 2.6. *If $h(n) := \omega(n) - \omega_Y(n)$, then we obtain*

$$\mathbb{P}\left(|h| \le (\log\log y)^{1/4}\right) = 1 - o(1),$$

where $\mathbb{P}$ denotes the probability value.

Proof. The first step is to find an estimate for $\mathbb{E}[h]$, and we have

$$\mathbb{E}[h] = \mathbb{E}\left[\omega(n) - \omega_Y(n)\right] = \mu_\omega(x,y) - \mu_{\omega_Y}(x,y).$$

Now by using Lemma 2.3 and 2.5, we get

$$\mathbb{E}[h] \ll (\log\log\log y)^{3/2}. \tag{15}$$

Finally we get our result by Markov's inequality and using (15)

$$\mathbb{P}\left(h \ge (\log\log y)^{1/4}\right) \le \frac{\mathbb{E}[\mathbb{h}]}{(\log\log y)^{1/4}} = o(1), \tag{16}$$

and the proof is complete. □

The above lemma and Remark 1 help us to observe that statement (4) can be expressed by the statement

$$\frac{1}{\Psi(x,y)}\#\{n\in S(x,y):\frac{\omega_Y(n)-\log\log y}{\sqrt{\log\log y}}\leq z\}\to\Phi(z)\quad\text{as}\quad y\to\infty, \tag{17}$$

and the above statement which will help us to prove our main theorem (Theorem 1.1) in the next section.

3. Proof of Theorem 1.1

We begin the proof by setting a family of independent random variables X_p on a probability space which satisfy

$$P(X_p=1)=\frac{\Psi(\frac{x}{p},y)}{\Psi(x,y)},\quad\text{and}\quad P(X_p=0)=1-\frac{\Psi(\frac{x}{p},y)}{\Psi(x,y)}. \tag{18}$$

And we define the partial sum S_Y as follows:

$$S_Y:=\sum_{p\leq Y}X_p,$$

where $Y=y^{1/\phi(y)}$. By X_p's properties in 18 and the estimate in (5) and (9), we deduce that $\mathbb{E}(S_Y)=\mathbb{V}(S_Y)=O\left(\log\log y\right)$ in the range $u=o(\log\log y)$, this means that $\omega_Y(n)$ and S_Y have approximately the same variance and average.

In the following lemma we get an upper bound for the difference of jth moments of ω_Y and S_Y, for $j=1,2,3,\cdots$.

Lemma 3.1. *If $u=o(\log\log y)$, then for any positive integer j we have*

$$A_j:=\mathbb{E}_{n\in S(x,y)}[\omega_Y(n)^j]-\mathbb{E}[S_Y^j]\ll_j\frac{(\log\log y)^j}{u(\log\log y)^{\sqrt{\log\log\log y}}}.$$

Proof. By the definition of ω_Y and S_Y, we have

$$\mathbb{E}[\omega_Y^j(n)]=\frac{1}{\Psi(x,y)}\sum_{p_1,\cdots,p_j\leq Y}\sum_{n\in S(x,y)}\mathbb{1}_{p_1|n}(n)\cdots\mathbb{1}_{p_j|n}(n),$$

and

$$\mathbb{E}[S_Y^j]=\sum_{p_1,\cdots,p_j\leq Y}\mathbb{E}\left[X_{p_1}\cdots X_{p_j}\right].$$

Without loss of generality, we assume that the p_i's are distinct. By expanding the difference of jth moments, we can express A_j using a new expression

$$\begin{aligned} A_j &= \sum_{p_1,\cdots,p_j \le y^{\frac{1}{\phi(y)}}} \left(\frac{1}{\Psi(x,y)} \sum_{n \in S(x,y)} \mathbb{1}_{p_1|n}(n) \cdots \mathbb{1}_{p_j|n}(n) - \mathbb{E}[X_{p_1} \cdots X_{p_j}] \right) \\ &= \sum_{p_1,\cdots,p_j \le y^{\frac{1}{\phi(y)}}} \left[\frac{\Psi(\frac{x}{p_1 \cdots p_j}, y)}{\Psi(x,y)} - \prod_{1 \le i \le j} \frac{\Psi(\frac{x}{p_i}, y)}{\Psi(x,y)} \right] \\ &= \sum_{p_1,\cdots,p_j \le y^{\frac{1}{\phi(y)}}} \left[\frac{\Psi(\frac{x}{p_1,\cdots,p_j}, y)}{\Psi(x,y)} - \prod_{1 \le i \le j} \frac{\Psi(\frac{x}{p_i}, y)}{\Psi(x,y)} \right]. \end{aligned} \tag{19}$$

By applying estimate (5) into the above equality, we obtain

$$\begin{aligned} A_j = & \sum_{p_1,\cdots,p_j < y^{1/\phi(y)}} \frac{1}{(p_1 \cdots p_j)^\alpha} \left\{ 1 + O\left(\frac{1}{u_y} + \frac{\log p_1 \cdots p_j}{\log x} \right) \right\} \\ & - \sum_{p_1,\cdots,p_j < y^{1/\phi(y)}} \frac{1}{(p_1 \cdots p_j)^\alpha} \prod_{i=1}^{j} \left\{ 1 + O\left(\frac{1}{u_y} + \frac{\log p_i}{\log x} \right) \right\}. \end{aligned}$$

The main terms in A_j are equal and will be eliminated. Therefore,

$$\begin{aligned} A_j &\ll \sum_{p_1,\cdots,p_j < y^{1/\phi(y)}} \frac{1}{(p_1 \cdots p_j)^\alpha} \left(\frac{1}{u_y} + \frac{\log p_1 \cdots p_j}{\log x} \right) \\ &\ll \sum_{p_1,\cdots,p_j < y^{1/\phi(y)}} \frac{1}{(p_1 \cdots p_j)^\alpha} \left(\frac{1}{u_y} + \frac{\log y}{\phi(y) \log x} \right). \end{aligned} \tag{20}$$

If $u = o(\log\log y)$, then $u_y \ge \frac{\log y}{\log\log\log y}$, so we can ignore the term $\frac{1}{u_y}$. Thus,

$$A_j \ll \sum_{p_1,\cdots,p_j < y^{1/\phi(y)}} \frac{1}{(p_1 \cdots p_j)^\alpha} \left(\frac{\log y}{\phi(y) \log x} \right).$$

We now use Lemma 2.5 and we get a new upper bound on A_j:

$$A_j \ll \frac{(\log\log y)^j}{u (\log\log y)^{\sqrt{\log\log\log y}}}. \tag{21}$$

□

Proof of Theorem 1.1. We start our proof by normalizing the random variable S_Y. Define

$$S := \frac{S_Y - \mu_{\omega_Y}(x,y)}{\sqrt{\sigma^2_{\omega_Y}(x,y)}}.$$

By recalling the Central Limit Theorem, one can say that S has a normal distribution $\Phi(x)$ since the X_p's are independent. We set

$$W := \frac{\omega_Y(n) - \mu_{\omega_Y}(x,y)}{\sqrt{\sigma^2_{\omega_Y}(x,y)}}.$$

By using the method of moments, we will show that the kth moments of W are very close to those corresponding to S, and they both converge to the kth moment of the normal distribution for every positive integer k. By the Multinomial Theorem, we get

$$\begin{aligned}\Delta^k &:= \mathbb{E}[(\omega_Y(n) - \mu_{\omega_Y}(x,y))^k] - \mathbb{E}[(S_Y - \mu_{\omega_Y(x,y)})^k] \\ &= \sum_{j=1}^{k} \binom{k}{j} (-\mu_{\omega_Y}(x,y))^{k-j} \left(\mathbb{E}[\omega_Y(n)^j] - \mathbb{E}[S_Y^j]\right).\end{aligned} \tag{22}$$

By combining the upper bound in (21) with (22), we arrive at an upper bound for Δ^k:

$$\begin{aligned}\Delta^k &\ll \frac{1}{(\log\log y)^{\sqrt{\log\log\log y}}} \sum_{j=1}^{k} \binom{k}{j} (-\mu_{\omega_Y}(x,y))^{k-j} (\log\log y)^j \\ &= \frac{1}{u(\log\log y)^{\sqrt{\log\log\log y}}} \left(\log\log y + \mu_{\omega_Y}(x,y)\right)^k.\end{aligned} \tag{23}$$

Now by using Lemma 2.3, we have

$$\Delta^k \ll \frac{(\log\log y)^k}{u(\log\log y)^{\sqrt{\log\log\log y}}}. \tag{24}$$

Thus,

$$\Delta^k \to 0 \quad \text{as} \quad x, y \to \infty.$$

We showed that the difference of kth moments goes to 0 for large values of y. By Remark 2, we conclude that two random variables S and W following the same distribution.

By Remark 4, the random variable S converges to a normal distribution. It remains to show that the moments of S are very close to those of the normal distribution. By recalling Remark 3, we need to prove that the moment $E[S^k]$ is bounded with respect to x and y when k increases. In fact, we will show that for each $k \in \mathbb{N}$:

$$\sup_{x,y} \left| \mathbb{E}\left(\frac{(S_Y - \mu_{\omega_Y}(x,y))^k}{(\sqrt{\sigma^2_{\omega_Y}(x,y)})^k}\right) \right| < \infty. \tag{25}$$

To complete the proof, we define a new family of independent random variables

$$Y_p = X_p - \frac{\Psi(x/p, y)}{\Psi(x, y)},$$

and we can reformulate the upper bound of the kth moment as follows:

$$\mathbb{E}\left((S_Y - \mu_{\omega_Y}(x, y))^k\right) = \sum_{j=1}^{k} {\sum}' \frac{k!}{k_1! \cdots k_j!} \sum_{p_1 \cdots p_j \leq y^{\frac{1}{\phi(y)}}} \mathbb{E}[Y_{p_1}^{k_1}] \cdots \mathbb{E}[Y_{p_j}^{k_j}], \qquad (26)$$

where ${\sum}'$ is over the j-tuple $(k_1, \cdots, k_j)$, $k_1, \cdots, k_j$ are positive integers, and $k_1 + \cdots + k_j = k$.
By the definition of $Y_p's$, we have $\mathbb{E}[Y_{p_j}] = 0$.
To avoid zero terms, we can assume that $k_i > 1$ for each $1 \leq i \leq j$. Also we have $|Y_p| \leq 1$. Thus,

$$\mathbb{E}[Y_p^{k_i}] \leq \mathbb{E}[Y_p^2], \quad \forall k_i > 2.$$

Therefore, the value of the inner sum in (26) is at most

$$\sum_{p_1 \cdots p_j \leq y^{\frac{1}{\phi(y)}}} \mathbb{E}[Y_{p_1}^{k_1}] \cdots \mathbb{E}[Y_{p_j}^{k_j}] \leq \left(\sum_{p \leq y^{\frac{1}{\phi(y)}}} \mathbb{E}[Y_p^2] \right)^j = \sigma^{2j}(x, y).$$

Each k_i is strictly greater than 1, and we have $k_1 + \cdots + k_j = k$. Therefore, $2j \leq k$ and this implies that

$$\mathbb{E}\left(\frac{(S_Y - \mu_{\omega_Y}(x, y))^k}{\left(\sqrt{\sigma^2_{\omega_Y}(x, y)}\right)^k} \right) \leq \sum_{j=1}^{k} {\sum}' \frac{k!}{k_1! \cdots k_j!},$$

from which (25) follows. We proved all necessary and sufficient conditions such that (17) and consequently (4) are true. □

Acknowledgement The author would like to thank Andrew Granville and Dimitris Koukoulopoulos for all their advice and encouragement as well as their valuable comments on the earlier version of the present paper. The author is also grateful to Adam Harper, Sary Drappeau and Oleksiy Klurman for helpful conversations.

References

[1] K. Alladi, An Erdős-Kac theorem for integers without large prime factors, *Acta Arith.* **49** (1987), 81-105.

[2] P. Billingsley, Patrick, On the central limit theorem for the prime divisor functions, *Amer. Math. Monthly* **76** (1969), 132-139.

[3] R. de la Bretèche and G. Tenenbaum, Entiers friables: inégalité de Turán-Kubilius et applications, *Invent. Math.* **159** (2005), 531-588.

[4] R. de la Bretèche and G. Tenenbaum, Propriétés statistiques des entiers friables, *Ramanujan J.* **9** (2005), 139-202.

[5] P. Erdös and M. Kac , The Gaussian law of errors in the theory of additive number theoretic functions, *Amer. J. Math.* **62** (1940), 738-742.

[6] W. Feller, *An introduction to probability theory and its applications, Vol. 1, 3rd ed*, John Wiley and Sons, 1968.

[7] A. Granville and K. Soundararajan, Sieving and the Erdős-Kac theorem, in *Equidistribution in number theory, an introduction*, **237** (2007), 15-27.

[8] D. Hensley, The distribution of $\Omega(n)$ among numbers with no large prime factors, *Analytic number theory and Diophantine proble (Stillwater, OK, 1984)* **70** (1983), 247-281.

[9] A. Hildebrand, On the number of prime factors of integers without large prime divisors, *J. Number Theory* **25** (1987), 81-106.

[10] A. Rényi and P. Turàn, On a theorem of Erdös-Kac, *Acta Arithmetica* **4** (1958), 71-84.

[11] G. Tenenbaum, On ultrafriable integers, *Q. J. Math.* **66** (2015), 333–351.