"Exercise" : try to prove the Erdős-Kac Theorem using Prop. B.4.4, to see that it does (seem to, at least) require different tools.

We will in fact prove the theorem using the method of moments. This is not the only approach but it is currently the simplest.

Th. (B.5.5 + B.5.6)

Let $(X_n)_{n\geq 1}$ be a sequence of $\mathbb{R}$-valued (or $\mathbb{C}$-valued) random variables such that the following conditions hold:

(1) for all $n \geq 1$ and all $k \geq 1$,
$$\mathbb{E}(|X_n|^k) < +\infty$$
(e.g., each X_n is bounded almost surely).

(2) for all $k \geq 1$ integer
$$\lim_{n \to \infty} \mathbb{E}(X_n^k) = \mathbb{E}(X^k)$$
for some other r.v. X s.t. $\mathbb{E}(|X|^k) < +\infty$.

(3) the series
$$\sum_{k\geq 0} \mathbb{E}(|X|^k) \frac{z^k}{k!}$$
has > 0 radius of convergence.

Then $X_n \xrightarrow{\text{law}} X$.

Conversely ~~furthermore~~ if (1) holds and $X_n \to X$ and

(4) for all $k \geq 1$, there exists $c_k \geq 0$ s.t.
$$\forall n \geq 1, \quad \mathbb{E}(|X_n|^k) \leq c_k$$
then (3) holds and so does (2).

A key feature of the method of moments from our point of view is that there is no independance assumption.

We now start the proof of the Erdős–Kac Theorem.

Step 1 – (Truncation)

One positive feature of the renormalization is that it allows us easily to reduce even further than $p \leq N$ the range of prime that matter.

Precisely, we define

$$Q = N^{1/(\log\log N)^{1/3}} = N^{1/\sigma_N^{1/3}}$$

$\left[\sigma_N = \sum_{p \leq N} \frac{1}{p}\right]$

$\left[0 < \frac{1}{3} < \frac{1}{2}\right]$

and

$$\tilde{\omega} : \Omega_N \longrightarrow \mathbb{R}$$
$$n \longmapsto \sum_{\substack{p \mid n \\ p \leq Q}} 1 = \sum_{p \leq Q} B_{p,N}(n)$$

Then observe that the number of primes $p \mid n$ "missing" from $\tilde{\omega}$ satisfies

$$\sum_{\substack{p \mid n \\ p > Q}} 1 \leq \sigma_N^{1/3}$$

$$\left(\text{since } \begin{array}{l} p_1 \cdots p_m \mid n \\ p_i > Q \end{array} \Rightarrow N^{m/\sigma_N^{1/3}} \leq N\right)$$

and hence

$$\frac{\tilde{\omega}}{\sigma_N^{1/2}} = \frac{\omega}{\sigma_N^{1/2}} + \underbrace{O\left(\frac{1}{\sigma_N^{1/6}}\right)}_{\text{uniformly}}$$

from which it is probabilistically elementary to deduce that if $\frac{\tilde{\omega}_N - \tilde{\sigma}_N}{\sqrt{\tilde{\sigma}_N}}$ converges in law, then so does

$$\frac{\omega_N - \sigma_N}{\sqrt{\sigma_N}} \qquad \text{(lemma B.5.3)}$$

[here we use the fact that

$$\begin{aligned}
\tilde{\sigma}_N &= \sum_{p \leq Q} \frac{1}{p} \\
&= \sum_{p \leq N^{1/\sigma_N^{1/3}}} \frac{1}{p} \\
&= \log\left(\log N^{1/\sigma_N^{1/3}}\right) + O(1) \\
&= \log\left(\frac{1}{\sigma_N^{1/3}} \log N\right) + O(1) \\
&= \log\log N - \frac{1}{3}\log \sigma_N + O(1) \\
&\underset{N \to \infty}{\sim} \sigma_N. \quad]
\end{aligned}$$

Finally as a last preparation, we write

$$\tilde{\omega}_{0,N} = \tilde{\omega}_N - \tilde{\sigma}_N = \sum_{p \leq Q} \left(B_{p,N} - \frac{1}{p}\right).$$

<u>Step 2.</u> (Computing the moments)

Let $k \geq 1$ be a fixed integer.

We have

$$\forall N, \quad \mathbb{E}_N\left(|\tilde{\omega}_{0,N}|^k\right) < +\infty$$

because these r.v. are bounded in fact.

So our goal is to prove that

$$\lim_{N \to \infty} \mathbb{E}_N\left(\tilde{\omega}_{0,N}^k\right) = \mathbb{E}\left(\mathcal{N}(0,1)^k\right) = \begin{cases} 0 & k \text{ odd} \\ k!/(k/2)!, & k \text{ even} \end{cases}$$

(We will do this *without* knowing beforehand the values of the *moments* $\mathbb{E}(\mathcal{N}(0,1)^k)$ of a standard Gaussian random variable.)

We compute directly

$$\mathbb{E}_N\left(\left(\frac{\tilde{\omega}_{0,N}}{\sigma_N^{1/2}}\right)^k\right)$$

$$= \frac{1}{\sigma_N^{k/2}} \mathbb{E}_N\left(\left(\sum_{p \le Q} \left(B_{p,N} - \frac{1}{p}\right)\right)^k\right)$$

$$= \frac{1}{\sigma_N^{k/2}} \sum_{\substack{p_1, \ldots, p_k \\ p_i \le Q}} \cdots \sum \mathbb{E}_N\left(\prod_{1 \le i \le p} \left(B_{p_i,N} - \frac{1}{p_i}\right)\right)$$

Fix $p_1, \ldots, p_k$. Then the r.v. on Ω_N

$$\prod \left(B_{p_i,N} - \frac{1}{p_i}\right)$$

is of the form

$$f(n \bmod q)$$

where

$$q = p_1 \cdots p_k \le Q^k$$

and

$$f(x) = \prod_{i=1}^{k} \left(\delta_{x \bmod p_i = 0} - \frac{1}{p_i}\right)$$

Since $|f| \le 1$ we have $\|f\|_1 \le q$, and hence by the basic theorem (1.3.1 in the script) we get

$$\mathbb{E}_N\left(\prod_{i=1}^{k}\left(B_{p_i} - \frac{1}{p_i}\right)\right) = \mathbb{E}_q(f) + \underbrace{O\left(\frac{q}{N}\right)}_{O\left(\frac{Q^k}{N}\right)}$$

Now the "main term" can be computed "backwards": we have

$$\mathbb{E}_q(f) = \mathbb{E}\left(\prod_{i=1}^{k} \left(B_{p_i} - \frac{1}{p} \right) \right)$$

where the (B_p) are independent Bernoulli r.v. with $\mathbb{P}(B_p = 1) = \frac{1}{p}$. ~~this follows~~ So, ?
by running the computation backwards, we obtain

$$\frac{1}{\sigma_N^{k/2}} \mathbb{E}_N\left(\tilde{\omega}_{0,N}^{k} \right) = \frac{1}{\sigma_N^{k/2}} \mathbb{E}\left(\left(\sum_{p \leq Q} \left(B_p - \frac{1}{p} \right) \right)^{k} \right)$$

$$(*) \qquad + O\left(\frac{1}{\sigma_N^{k/2}} \frac{Q^{2k}}{N} \right)$$

The choice of Q is such that for any <u>fixed</u> $k \geq 1$, we have

$$\frac{Q^{2k}}{N} \xrightarrow[N \to \infty]{} 0 .$$

<u>Step 3.</u> (CLT)

The basic formula $(*)$ shows how the statistical properties of ω are dictated by those of the "expected" heuristic model

$$\sum_{p \leq Q} \left(B_p - \frac{1}{p} \right) = \sum_{p \leq N} \left(B_p - \frac{1}{p} \right) \cancel{+ O(\log\log}$$

$$+ O\left(\sigma_N^{1/3} \right)$$

as before.

Now we just need to conclude with two probabilistic facts.

(1) we have $\frac{1}{\sqrt{\sigma_N}} \sum_{p \leq N} \left(B_p - \frac{1}{p}\right) \xrightarrow[\text{law}]{} \mathcal{N}(0,1)$

(B. 7. 1)

(2) we have furthermore for any $k \geq 1$

$$\lim_{N \to +\infty} \mathbb{E}\left(\left(\frac{1}{\sqrt{\sigma_N}} \sum_{p \leq N} \left(B_p - \frac{1}{p}\right)\right)^k\right) = \mathbb{E}\left(\mathcal{N}(0,1)^k\right).$$

This follows from Th. B. 5. 6 provided we can prove the "uniform integrability"

$$\forall N \geq 1, \quad \mathbb{E}\left(\left(\frac{1}{\sqrt{\sigma_N}} \sum_{p \leq N} \left(B_p - \frac{1}{p}\right)\right)^k\right) \leq c_k$$

We do this as follows. Write

$$\tilde{B}_p = B_p - \frac{1}{p}$$

and $X_N = \frac{1}{\sqrt{\sigma_N}} \sum_{p \leq N} \tilde{B}_p$

Note that $\mathbb{E}(\tilde{B}_p) = 0$, $|\tilde{B}_p| \leq 1$.

~~and~~

We will show that there exists in fact a constant $C \geq 0$ s.t.

$$\begin{cases} \mathbb{E}(e^{X_N}) \leq C \\ \mathbb{E}(e^{-X_N}) \leq C \end{cases}$$

Since $|x|^k \leq c_k (e^x + e^{-x})$ for all $x \in \mathbb{R}$, this implies the result.

Now let $t = \pm 1$ (or $|t| \leq 1$).

By independance, we have

$$\mathbb{E}\left(e^{tX_N}\right)$$

$$= \prod_{p \leq N} \mathbb{E}\left(\exp\left(\underbrace{\frac{t\tilde{B}_p}{\sqrt{\sigma_N}}}_{\leq 1 \text{ if } N \text{ large enough}}\right)\right)$$

$$\leq \prod_{p \leq N} \mathbb{E}\left(1 + \frac{t\tilde{B}_p}{\sqrt{\sigma_N}} + \frac{t^2 \tilde{B}_p^2}{\sigma_N}\right)$$

$$= \prod_{p \leq N} \left(1 + \frac{t^2}{\sigma_N} \mathbb{V}(\tilde{B}_p)\right)$$

$$\leq \prod_{p \leq N} \exp\left(\frac{t^2}{\sigma_N} \mathbb{V}(\tilde{B}_p)\right)$$

$$= \exp\left(\frac{t^2}{\sigma_N} \sum_{p \leq N} \mathbb{V}(\tilde{B}_p)\right)$$

$$\leq \exp\left(t^2\right).$$

This concludes the proof.