Euler Convergence: Probabilistic Considerations

GIORGIO METAFUNE
DIEGO PALLARA
CARLO SEMPI
Dipartimento di Matematica
"Ennio De Giorgi"
Università di Lecce
Lecce, Italy 73100

In mathematical analysis and its applications, the need sometimes arises to generalize the concept of limit of a sequence (or of sum of a series) in order to include cases in which the sequence (or the series) does not converge in the ordinary sense. One of the several methods devised bears Euler's name.

A sequence of real numbers $\{x_n\}$ is said to converge to x *in the sense of Euler* (or, usually, to be *Euler-convergent*) if there exists s in the open interval $(0,1)$ such that

$$\lim_{n\to\infty}\sum_{j=0}^{n}\binom{n}{j}s^j(1-s)^{n-j}x_j = x.$$

The first key theorem on Euler convergence is as follows:

THEOREM 1. *If a sequence $\{x_n\}$ of real numbers converges to $x\in\mathbb{R}$ then it converges to the same limit in the sense of Euler for every $s\in(0,1)$.*

Euler convergence is used in a natural way in the study of the asymptotic behavior of cyclic Markov chains (see **[5]**, Chapter V). While teaching Markov chains using the approach of **[5]**, we noticed that Theorem 1 could be proved in an elementary manner by relying only on the Chebyshev inequality, which appears in introductory courses in probability. Our proof provides a good chance to practice probabilistic reasoning. For a purely analytical proof of Theorem 1, one is often (e.g., in **[3]**) referred to **[4]**. But Hardy's book is hard reading for an undergraduate—even more so since it deals with series rather than with sequences.

A sequence of real numbers may be Euler-convergent without being convergent. Consider, for instance the sequence $\{x_n\}$ with $x_{2j}=1$, $x_{2j+1}=0$ $(j\ge 0)$, which does not converge. As for Euler convergence, notice that

$$\sum_{j=0}^{n}\binom{n}{j}s^j(1-s)^{n-j}x_j = \sum_{\substack{j=0\\ j \text{ even}}}^{n}\binom{n}{j}s^j(1-s)^{n-j};$$

this latter sum represents the probability that a binomial random variable S_n takes an even value. Here $S_n=\sum_{j=1}^{n}X_j$, where the X_j's are Bernoulli random variables with $P(X_n=1)=s$. By the binomial theorem,

$$\sum_{\substack{j=0\\ j \text{ even}}}^{n}\binom{n}{j}s^j(1-s)^{n-j}+\sum_{\substack{j=0\\ j \text{ odd}}}^{n}\binom{n}{j}s^j(1-s)^{n-j}=(s+1-s)^n=1,$$

and

$$\sum_{\substack{j=0\\ j \text{ even}}}^{n} \binom{n}{j}(-s)^j(1-s)^{n-j} + \sum_{\substack{j=0\\ j \text{ odd}}}^{n} \binom{n}{j}(-s)^j(1-s)^{n-j} = (1-2s)^n.$$

Adding the last two equations yields

$$P\left(\bigcup_{j \text{ even}} \{S_n = j\}\right) = \sum_{\substack{j=0\\ j \text{ even}}}^{n} \binom{n}{j} s^j(1-s)^{n-j} = \frac{1+(1-2s)^n}{2};$$

this tends to $1/2$ as n tends to ∞.

Now we prove that Euler convergence is implied by ordinary convergence.

Proof of Theorem 1. Since

$$\sum_{j=0}^{n} \binom{n}{j} s^j(1-s)^{n-j} = (s+1-s)^n = 1,$$

it suffices to consider sequences that converge to zero. Let $\{x_n\}$ be such a sequence. For a fixed $\varepsilon \in (0, s)$ there exists $\nu \in \mathbb{N}$ such that $|x_n| < \varepsilon$ for every $n \geq \nu$. Now let $\lambda := \max\{|x_j| : j = 0, 1, \ldots\}$ Then, for $n > \nu$, we have

$$\begin{aligned}
\left|\sum_{j=0}^{n} \binom{n}{j} s^j(1-s)^{n-j} x_j\right| &\leq \left|\sum_{j=0}^{\nu-1} \binom{n}{j} s^j(1-s)^{n-j} x_j\right| + \left|\sum_{j=\nu}^{n} \binom{n}{j} s^j(1-s)^{n-j} x_j\right| \\
&< \lambda \sum_{j=0}^{\nu-1} \binom{n}{j} s^j(1-s)^{n-j} + \varepsilon \sum_{j=\nu}^{n} \binom{n}{j} s^j(1-s)^{n-j} \\
&< \lambda \sum_{j=0}^{\nu-1} \binom{n}{j} s^j(1-s)^{n-j} + \varepsilon.
\end{aligned}$$

Notice that if $n > (\nu-1)/(s-\varepsilon)$ then $ns - (\nu-1) > n\varepsilon$ and, *a fortiori*, $ns - j > n\varepsilon$ for $j = 0, 1, \ldots, \nu-1$. Finally, observe that

$$\sum_{j=0}^{\nu-1} \binom{n}{j} s^j(1-s)^{n-j} = P\left(\bigcup_{j=0}^{\nu-1} \{S_n = j\}\right) \leq P(|S_n - ns| \geq n\varepsilon).$$

Now, since the variance of the binomial distribution is equal to $ns(1-s)$ if $s \in (0,1)$ is the probability of success, Chebyshev's inequality yields

$$\sum_{j=0}^{\nu-1} \binom{n}{j} s^j(1-s)^{n-j} \leq \frac{ns(1-s)}{n^2\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \to 0.$$

Therefore,

$$\left|\sum_{j=0}^{n} \binom{n}{j} s^j(1-s)^{n-j} x_j\right|$$

tends to zero as $n \to \infty$.

Alternatively, in a more analytical vein, one could avoid recourse to the Chebysbev inequality by considering the sum

$$\sum_{j=0}^{\nu-1} \binom{n}{j} s^j (1-s)^{n-j} \tag{1}$$

and noting that the number of terms (ν) is fixed, and that the binomial coefficient tends to infinity like n^j while the factor $(1-s)^{n-j}$ tends to zero exponentially. Thus each term in the sum (1), and hence the whole sum, tends to zero. Even this second approach has a probabilistic meaning: it is well-known (see, e.g., [**7**; Section 3.3]) that the terms of the binomial distribution increase in j from 0 to $(n+1)s$ and decrease when j runs from $(n+1)s+1$ to n. Since ν is fixed , the sum (1) represents, when n goes to infinity, the ever-decreasing probability of the tail of the binomial distribution. □

The first proof given above is an adaptation of the argument used by Bernstein [**1**] in his proof of the Weierstrass theorem on uniform approximation by polynomials of all functions that are continuous on a closed interval. A modern presentation can be found in the exercises of [**6**] or in [**2**].

REFERENCES

1. S. Bernstein, Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités, *Soob Charkov Mat. Obs.* 13 (1912), 1–2.
2. Y. S. Chow and H. Teicher, *Probability Theory, Independence, Interchangeability, Martingales*, Springer-Verlag, New York–Heidelberg–Berlin, 1978.
3. *Encylopaedia of Mathematics, Volume* 3, Kluwer, Dordrecht–Boston–London, 1989.
4. G. H. Hardy, *Divergent Series*, Clarendon, Oxford, UK, 1949.
5. J. G. Kemeny and J. L. Snell, *Finite Markov Chains*, Van Nostrand, New York, NY, 1960 (reprinted, Springer-Verlag, New York, NY, 1976).
6. M. Loève, *Probability Theory*, Van Norstrand, New York, NY, 1963 (reprinted in two volumes, Springer-Verlag, New York–Heidelberg–Berlin, 1997).
7. E. Parzen, Modern Probability and Its Applications, Wiley, New York and London, 1960.