Computational Sciences and Engineering 5(2) (2025) 287-293

University of Guilan

Computational Sciences and Engineering

journal homepage: https://cse.guilan.ac.ir/

A Probabilistic Approach to the Interpolation Error

Sadia Kumail Awad [a], Ali Tavakoli [b,*]

[a] Polytechnic College/Al-Qadisiyah, Al-Furat Al-Awsat Technical University, Iraq
[b] Applied mathematics department, Faculty of mathematical sciences, University of Mazandaran, Babolsar, Iran

ARTICLE INFO

Article history:
Received 22 February 2026
Received in revised form 12 April 2026
Accepted 12 April 2026
Available online 12 April 2026

Keywords:
Polynomial interpolation
Chebyshev nodes
Trapezoidal rule

ABSTRACT

Polynomial interpolation is a fundamental tool in numerical analysis, with its accuracy classically characterized by the Lagrange remainder term. This term involves an unknown mean value point $\xi \in [a, b]$, which depends on x and the function f. Consequently, while the formula provides an exact error representation, its practical utility for a priori error estimation is limited, as determining a sharp, computable bound for high-order derivatives is often challenging. This paper introduces a novel probabilistic framework to address this longstanding limitation. Instead of treating ξ as an unknown deterministic value, we model its location probabilistically. By considering the interpolation nodes as random variables or analyzing the distribution of ξ for a fixed x, we derive a statistical estimate for its expected value. This approach allows us to propose a specific, computable value for ξ that provides a highly accurate estimate of the actual interpolation error. The theoretical findings are substantiated with several numerical examples. These experiments demonstrate that our probabilistic estimate of the remainder term consistently aligns with the true error, offering a practical and powerful alternative to traditional worst-case error bounds. This method provides a new perspective on error analysis in approximation theory, bridging deterministic numerical methods with probabilistic techniques.

1. Introduction

The error term of a numerical scheme determines its effectiveness whereby, for example, Iserles showed the Gauss-Christoffel quadrature rule works well for integration of smooth functions, however for high oscillatory functions, it is a completely useless estimation (see for instance [1,2]). Furthermore, by choosing the suitable upper bounds on the error of trapezoidal integration, Trefethen and Weideman [3] verified the exponential order of the error. Finding the numerical

* Corresponding author.
E-mail addresses: a.tavakoli@umz.ac.ir (A. Tavakoli)

https://doi.org/10.22124/cse.2026.33090.1163

errors of interpolation (see for instance [4]) as well as the explicit upper bound for the interpolation error (see for instances [5, 6]) is still of interest to many researchers . For the error of polynomial interpolation in the one-dimensional case, a famous upper bound is seen in many textbooks (see for instance [7]). In this paper, we discuss on the desirability of this upper bound. In fact, the sharpness of this upper bound is discussed here.

Let f be a function defined on $[a,b]$ and $x_0, x_1, \cdots, x_n$ be $n+1$ distinct nodal points such that

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b. \quad \textbf{(1)}$$

Suppose P_n is the interpolating polynomial of f, at most of degree n, that interpolates the function f at the above nodal points. If $f \in C^{n+1}([a,b])$, it is well known that the interpolation error of f by P_n satisfies as

$$E(\bar{x}) = f(\bar{x}) - P_n(\bar{x}) = \frac{\Psi(\bar{x}) f^{(n+1)}(\xi)}{(n+1)!}, \quad \textbf{(2)}$$

$\bar{x} \in [a,b]$,

where $\Psi(x) = (x - x_0) \cdots (x - x_n)$ and $\xi \in (a,b)$ (see [7]).

The function $f^{(n+1)}$ is bounded, therefore, there exists $M > 0$ such that

$$\| f^{(n+1)} \|_\infty \leq M, \quad \textbf{(3)}$$

then

$$\| E \|_\infty \leq \frac{M \| \Psi \|_\infty}{(n+1)!}. \quad \textbf{(4)}$$

The question is that:

What is the use of the upper bound given in ***Eq. (4)***?

If $\| \Psi \|_\infty$ is small enough, it clearly depends on the value of M. For instance, let T_{n+1} be the Chebyshev polynomial of degree $n+1$ on $[a,b] = [-1,1]$. Then,

$$\min_{x_i} \| \Psi \|_\infty \quad \textbf{(5)}$$

is reached on the roots of Chebyshev polynomial of degree $n+1$. In this case, $\Psi(x) = 2^{-n} T_{n+1}(x)$ and so

$$\| \Psi \|_\infty = 2^{-n}, \quad \textbf{(6)}$$

which is an enough small value for enough large n.

In order to clarify the importance of the preceding question, we state an example:

Example 1 Let $\mathrm{f(x)} = \mathrm{e}^{10\mathrm{x}}$ be defined on [-1,1]. How many Chebyshev points should be chosen such that the error of polynomial interpolation on these points is at most 10^{-9}?

Answer: The relations

$$\| (e^{10x})^{(n+1)} \|_\infty = \| 10^{n+1} e^{10x} \|_\infty = 10^{n+1} e^{10} \quad \textbf{(7)}$$

and ***Eq. (3)*** imply that

$$\| E \|_\infty = \| f - P_n \|_\infty \quad \leq \quad \frac{2^{-n} 10^{n+1} e^{10}}{(n+1)!}. \quad \textbf{(8)}$$

It is seen that the smallest value of n such that $\frac{2^{-n}10^{n+1}e^{10}}{(n+1)!} \leq 10^{-9}$ is $n = 32$, while it is computable that $\| f - P_{27} \|_{\infty} \approx 7.5e - 10 < 10^{-9}$. Thus, the upper bound of (4) seems not to be suitable!

2. Probability theory

We start this section by a fundamental question:

How to find an appropriate upper bound of the error?

To respond to this question, we again note to ***Eq. (2)***. If we discover the place of ξ for any $x \in [a, b]$, specifically for $arg(\| E \|_{\infty})$, then by ***Eq. (6)***, $\| E \|_{\infty}$ will be exactly computed. We rewrite the question in another form:

How to find ξ in (a, b)?

Our aim is to response this question by probability theory. What is the probability of being a random value $\bar{x} \in (a, b)$ equal to ξ? We know that the answer is 0, because (a, b) is a continuous interval. Therefore, we change the question as follows: What is the probability of ξ falling in the subinterval $(x_i, x_{i+1}]$ for $i = 0,1,\cdots,n-1$?

To answer this question, we review the proof of the error for polynomial interpolation given in (1) (see for instance [7]):

Proof. If $\bar{\boldsymbol{x}} = \boldsymbol{x_i}$, there is nothing to show. Let $\bar{\boldsymbol{x}} \neq \boldsymbol{x_i}$ for $\boldsymbol{i = 0, 1, \cdots, n}$.

We can find a constant K such that the function $F(x) = f(x) - P_n(x) - K\Psi(x)$ vanishes for $x = \bar{x}$, i.e.,

$$F(\bar{x}) = 0. \tag{9}$$

Consequently, F has at least the $n + 2$ zeros $x_0, \cdots, x_n, \bar{x}$ in $[a, b]$. By Rolle's theorem, applied repeatedly, F' has at least $n + 1$ zeros in the above interval, F'' at least n zeros, and finally $F^{(n+1)}$ at least one zero $\xi \in (a, b)$.

Since $P^{(n+1)} \equiv 0$,

$$F^{(n+1)}(\xi) = f^{(n+1)}(\xi) - K(n+1)! = 0, \tag{10}$$

or

$$K = \frac{f^{(n+1)}(\xi)}{(n+1)!}. \tag{11}$$

This proves the theorem. ■

Now, we review the above proof of interpolation error and try to answer the last question. For simplicity, we think of the interval $[x_0, x_3]$ such that the nodal points are x_0, x_1, x_2 and x_3. Let the function F be the same as given in the proof of interpolation error. Then,

$$F(x_0) = F(x_1) = F(x_2) = F(x_3) = 0. \tag{12}$$

For simplicity, we left the point $\bar{x}$. According to the Rolle's theorem, there exist the constants t_1, t_2 and t_3 such that $F'(t_1) = F'(t_2) = F'(t_3) = 0$ where

$$x_0 < t_1 < x_1 < t_2 < x_2 < t_3 < x_3.$$

Repeating Rolle's theorem, there exist the constants e_1 and e_2 such that $F''(e_1) = F''(e_2) = 0$ in which $t_1 < e_1 < t_2 < e_2 < t_3$. Finally, one more time applying Rolle's theorem implies that $F'''(\xi) = 0$ for some $\xi \in (e_1, e_2)$.

All possible events of e_1, e_2 and ξ are as follows:

$$\begin{array}{llll}
(1) & e_1 \in (x_0, x_1), & e_2 \in (x_1, x_2), & \xi \in (x_0, x_1); \\
(2) & e_1 \in (x_0, x_1), & e_2 \in (x_1, x_2), & \xi \in (x_1, x_2); \\
(3) & e_1 \in (x_0, x_1), & e_2 \in (x_2, x_3), & \xi \in (x_0, x_1); \\
(4) & e_1 \in (x_0, x_1), & e_2 \in (x_2, x_3), & \xi \in (x_1, x_2); \\
(5) & e_1 \in (x_0, x_1), & e_2 \in (x_2, x_3), & \xi \in (x_2, x_3); \\
(6) & e_1 \in (x_1, x_2), & e_2 \in (x_1, x_2), & \xi \in (x_1, x_2); \\
(7) & e_1 \in (x_1, x_2), & e_2 \in (x_2, x_3), & \xi \in (x_1, x_2); \\
(8) & e_1 \in (x_1, x_2), & e_2 \in (x_2, x_3), & \xi \in (x_2, x_3).
\end{array} \quad \textbf{(13)}$$

Now, assuming a uniform distribution, the probability of ξ falling in each subinterval is easily computable (see ***Table 1***).

Table 1. The probability of ξ falling in each subinterval of (x_0, x_3)

Subinterval	(x_0, x_1)	(x_1, x_2)	(x_2, x_3)
Probability	2/8	4/8	2/8

Now, we are ready to state a general formula.

Theorem 1 Let F be an enough smooth function such that

$$\begin{aligned} &F(x_0) = F(x_1) = \cdots = F(x_n) = 0, \\ &x_0 < x_1 < \cdots < x_n, \qquad n \geq 3. \end{aligned} \quad \textbf{(14)}$$

Assuming a uniform distribution, the probability of ξ falling in the interval (x_{i-1}, x_i) for $i = 1,2, \ldots, n$ is

$$\begin{aligned} &P_i = \frac{O_i}{\sum_{j=1}^{\lceil n/2 \rceil} O_{2j-1} + \sum_{j=1}^{\lfloor n/2 \rfloor} O_{2j}}, \\ &O_{2j-1} = 2(j-1)(n-2(j-1)) + n - 2j, \\ &O_{2j} = 2(j-1)(n-(2j-1)) + 2(n-2j). \end{aligned} \quad \textbf{(15)}$$

Proof. We prove the theorem by induction on n. For $n = 3$, we have already verified it. Now, assume that the theorem is true for up to n points. It is readily seen that if one point is added, i more selection are added to O_i. This implies,

$$\begin{aligned} O_{2j-1} &= 2(j-1)\big(n-2(j-1)\big) + n - 2j + (2j-1) = 2(j-1)\big(n-2(j-1)\big) \\ &\quad + n - 2j + (2j-1) + 2(j-1) + 1 \\ &= 2(j-1)((n+1)-2(j-1)) + (n+1) - 2j, \\ O_{2j} &= 2(j-1)(n-(2j-1)) + 2(n-2j) + 2j \\ &= 2(j-1)(n-(2j-1)) + 2(n-2j) + 2(j-1) + 2j \\ &= 2(j-1)((n+1)-(2j-1)) + 2((n+1)-2j). \end{aligned} \quad \textbf{(16)}$$

This completes the proof.

Remark 1 *In the proof of interpolation error, there is a point* $\bar{x}$ *such that* $F(\bar{x}) = 0$. *Here, for simplicity, we removed this point and so we repeat Rolle's theorem* n *times instead of* $n + 1$ *times!*

Although the value of ξ depends on $\bar{x}$, the more n increases, the less probability P_i will change.

Corollary 1 *The maximum probability occurs in the middle subinterval(s).*

Proof. By ***Eq. (15)***, the denominator of P_i is fixed and so

$$\underset{i}{\operatorname{argmax}} P_i = \underset{i}{\operatorname{argmax}} O_i.$$

Let $i = 2j - 1$ and $f(j) := O_{2j-1} = -4j^2 + (2n + 6)j - (n + 4)$. Taking $df/dj = 0$ gives $j = (n + 3)/4$ and so $2j - 1 = (n + 1)/2$.

Again, let $i = 2j$ and $f(j) := O_{2j} = -4j^2 + 2(n + 1)j - 2$. Taking $df/dj = 0$ gives $j = (n + 1)/4$ and so $2j = (n + 1)/2$.

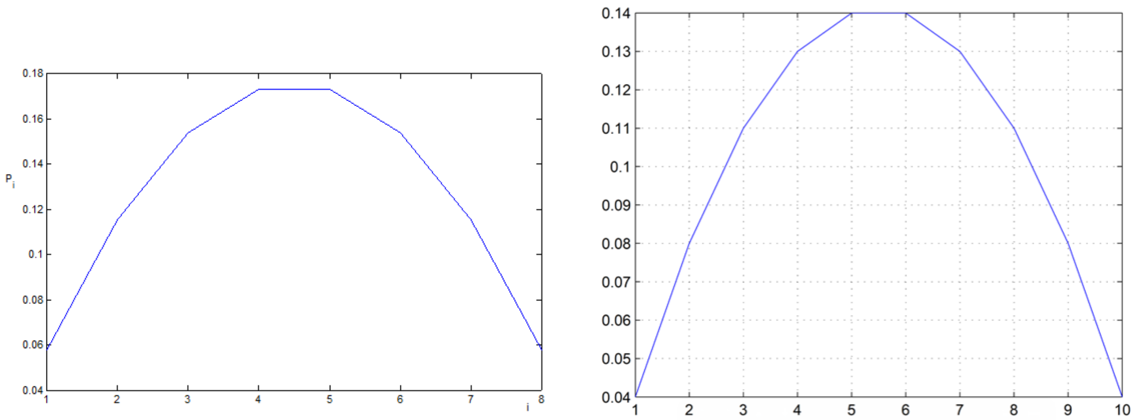
Special case: Let $n = 2m$ be an even number. Then, by Theorem 1,

$$P_1 = P_n = \frac{2m - 2}{\frac{2}{3}m(m - 1)(2m + 5)} = \frac{3}{m(2m + 5)},$$

and

$$P_m = \frac{(m - 1)(m + 2)}{\frac{2}{3}m(m - 1)(2m + 5)} = \frac{3(m + 2)}{2m(2m + 5)}.$$

It is observed that if m tends to infinity, as we expect, the probability P_i vanishes. ***Figure 1*** shows the probability P_i for $n = 8$ and $n = 10$. As is seen, the probability of event the middle subinterval(s) occurring is more than the



(a) (b)

Figure 1. The probability of the events each subinterval occurring for (a) $n = 8$ and (b) $n = 10$.

other subintervals. Moreover, for $n = 8$, the probability of $\xi \in (0,8)$ falling in the interval (2,6) is almost 65 percent and for $n = 10$, the probability of $\xi \in (0,10)$ falling in the interval (3,7) is 54 percent.

Finally, for the preceding example, if we take

$$\xi \in (-1 + 13 \times \frac{2}{27}, -1 + 14 \times \frac{2}{27})$$

(the middle interval of $[-1,1]$ for 28 nodal points), then

$$\frac{2^{-27} 10^{27+1} e^{-1+14\times\frac{2}{27}}}{(27+1)!} \leq 10^{-9}$$

In other words, the error bound of E satisfies for $n = 27$.

3. An application

In this section, we present an application of the preceding theorem. This is related to the presenting a modification of trapezoidal rule for estimation of a definite integral.

Let $f \in \mathcal{C}^2[a,b]$ and $x_0, x_1, \cdots, x_n$ be $n+1$ distinct nodal points such that

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b,$$

and

$$x_i - x_{i-1} = h, \qquad i = 1,2,\cdots,n.$$

The error of trapezoidal rule ($T(h)$) for integration is given by (Kincaid and Cheney (1990))

$$\int_a^b f(x)dx - T(h) = -\frac{b-a}{12} h^2 f''(\xi), \qquad a < \xi < b. \tag{17}$$

In the proof of this error, it is shown that

$$f''(\xi) = \frac{f''(\xi_1) + f''(\xi_2) + \cdots + f''(\xi_n)}{n}, \tag{18}$$

where $x_{i-1} < \xi_i < x_i$ for $i = 1, \cdots, n$. The value ξ_i is appeared from the interpolation error of f by trapezoidal rule in the subinterval $[x_{i-1}, x_i]$, i.e.,

$$\int_{x_{i-1}}^{x_i} f(x)dx - \frac{h}{2}(x_{i-1} + x_i) = \int_{x_{i-1}}^{x_i} \frac{(x - x_{i-1})(x - x_i)}{2!} f''(\eta(x))dx = -\frac{h^3}{12} f''(\xi_i). \tag{19}$$

Now, by Corollary 1, the probability of ξ_i falling in the middle subintervals of $[x_{i-1}, x_i]$ is maximum. Then, taking $m_i := \frac{x_{i-1}+x_i}{2}$, the value of f'' in ***Eq. (18)*** can be approximated by

$$f''(\xi) \approx \frac{f''(m_1) + f''(m_2) + \cdots + f''(m_n)}{n} \tag{20}$$

A standard approximation of f'' is given by

$$f''(m_i) \approx \frac{f(x_{i-1}) - 2f(m_i) + f(x_i)}{h_1^2}, \quad i = 1, \cdots, n \tag{21}$$

where $h_1 = h/2$. Substituting f'' from ***Eq. (21)*** into ***Eq. (17)*** yields:

$$\int_a^b f(x)dx \approx \frac{h_1}{3}(f(x_0) + 4\sum_{i=1}^{n} f(m_i) + 2\sum_{i=1}^{n} f(x_i) + f(x_n)), \quad (22)$$

that is the same Simpson's rule!

4. Conclusion

The error formula for polynomial interpolation of an enough smooth function $f: [a, b] \to R$ with $n + 1$ distinct nodal points x_i such that

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b,$$

depends on $f^{(n+1)}(\xi)$, where $\xi \in (a, b)$. In this work, by probability theory, we extracted a closed formula to compute the probability of ξ falling in (x_i, x_{i+1}), for $i = 0,1, \cdots, n - 1$. We considered a uniform distribution that means the value of per subinterval is the same. We have also presented an application of this probability theory.

References

[1] Iserles, A. (2004). On the numerical quadrature of highly-oscillating integrals I: Fourier transforms. *IMA Journal of Numerical Analysis*, *24*(3), 365-391.
[2] Iserles, A. (2005). On the numerical quadrature of highly-oscillating integrals II: Irregular oscillators. *IMA journal of numerical analysis*, *25*(1), 25-44.
[3] Trefethen, L. N., & Weideman, J. A. C. (2014). The exponentially convergent trapezoidal rule. *SIAM review*, *56*(3), 385-458.
[4] Kaltofen, E. L. (2022, July). Sparse polynomial Hermite interpolation. In *Proceedings of the 2022 International Symposium on Symbolic and Algebraic Computation* (pp. 469-478).
[5] Broniatowski, M., Celant, G., Di Battista, M., & Leoni-Aubin, S. (2012). Upper bounds for the error in some interpolation and extrapolation designs. *Communications in Statistics-Theory and Methods*, *41*(16-17), 3002-3019.
[6] Liu, X., & You, C. G. (2018). Explicit bound for quadratic Lagrange interpolation constant on triangular finite elements. *Applied Mathematics and Computation*, *319*, 693-701.
[7] Stoer, J., R. Bulirsch. 2002. *Introduction to Numerical Analysis*. 3rd edition. Springer.