

Article

Not peer-reviewed version

Practical Method for $\log[\operatorname{erfc}(a)]$ Approximation in the Extreme Tail

[Kazuharu Misawa](#) *

Posted Date: 9 March 2026

doi: 10.20944/preprints202603.0686.v1

Keywords: complementary error function (erfc); rare-event probability; numerical stability; geometric approximation methods



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Practical Method for $\log[\operatorname{erfc}(a)]$ Approximation in the Extreme Tail

Kazuharu Misawa ^{1,2,*}

¹ Yokohama City University Graduate School of Medicine, 22-2 Seto, Kanazawa-ku, Yokohama 236-0027, Japan; kazu_misawa@hotmail.com

² RIKEN Center for Advanced Intelligence Project (AIP), Nihonbashi 1-chome Mitsui Building, 15th floor, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

Abstract

Accurate evaluation of extremely small Gaussian tail probabilities is essential in statistical meta-analyses, in which large z -scores (often exceeding 8 or 9) must be converted into p -values. Meanwhile, direct numerical integration of complementary error function $\operatorname{erfc}(a)$ suffers from severe underflow in floating-point arithmetic. In this paper, a simple and robust approximation scheme for $\log[\operatorname{erfc}(a)]$ is proposed based on a geometric tangent construction. This approach yields explicit lower and upper bounds, closed-form asymptotic expansions up to order a^{-8} , and numerically stable formulas suitable for implementation in statistical software. Numerical comparisons demonstrate that the lower and upper bounds become extremely tight for $a \geq 6$, making the proposed method practical for large-scale meta-analytic computations.

Keywords: complementary error function (erfc); rare-event probability; numerical stability; geometric approximation methods

1. Introduction

Meta-analyses frequently combine results from multiple studies by summing z -scores, which are assumed to be approximately normally distributed. The sum distribution is also normal. Therefore, evaluating the combined p -value requires evaluating the Gaussian tail probabilities of the following form for large a :

$$\operatorname{erfc}(a) = \int_a^{\infty} \exp[-x^2] dx.$$

For $a \geq 8$, direct computation is numerically unstable: extremely small values underflow to zero in standard double-precision arithmetic.

Existing approximation techniques for the complementary error function $\operatorname{erfc}(a)$ are almost exclusively based on analytic minimax polynomial approximations, rational-function approximations, Chebyshev-series expansions, or Padé-type expansions, as extensively documented in the NIST Digital Library of Mathematical Functions [1-3]. To our knowledge, no prior work has introduced a geometric construction—such as a tangent-based area argument—to derive explicit upper and lower bounds for erfc or $\log[\operatorname{erfc}(a)]$ in the extreme tail.

The main purpose of this work was to construct a method that (i) avoids numerical underflow; (ii) provides explicit lower and upper bounds; (iii) admits asymptotic expansions to arbitrary order; and (iv) is easy to implement in practical statistical workflows.

2. Methods

2.1. Geometric Motivation

Consider a positive, decreasing function $h = f(x)$, and point $A = (a, h)$ on its graph. Let $s = f'(a)$ denote the slope, and let l be the horizontal distance from foot O of the vertical through A to x -intercept P of the tangent line at A . Then,

$$l = \frac{h}{|f'(a)|}.$$

The area of triangle AOP is

$$\text{Area}(AOP) = \frac{hl}{2} = \frac{f(a)^2}{2|f'(a)|}.$$

If f decays approximately exponentially, then the integral $\int_a^\infty f(x)dx$ is comparable to rectangle $h \times l$, whose area is $f(a)^2/|f'(a)|$.

At a point of tangency $x = a$ on the exponential curve $y = b \exp[-cx]$, the tangent line and the perpendicular to the x -axis from a right triangle (Figure 1). The area of this triangle is exactly half of tail integral

$$\int_a^\infty b \exp[-cx] dx$$

y

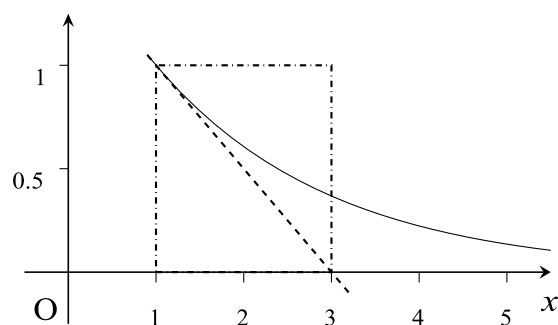


Figure 1. Geometric interpretation of the exponential tail. At the tangency point of the exponential curve, the tangent line and the perpendicular dropped to the x -axis form a right triangle whose area equals one-half of the integral of the exponential function beyond the tangency point.

2.2. Gaussian Tail and Tangent-Matched Lower Bound

This geometric principle is applied to the Gaussian function. The triangle area provides approximately half of this value, suggesting a natural upper and lower bound pair. For the tangency point chosen, the Gaussian tail $\text{erfc}(x)$ (dotted line) lies between the exponential decay curve (solid line) and the area of the right triangle generated by the tangent line (dashed line), demonstrating that the Gaussian tail integral is geometrically sandwiched between two quantities.

Let

$$f(x) = \exp[-x^2],$$

As $f'(a) = -2a \exp[-a^2]$, the rectangle bound gives

$$p_{\text{upper}}(a) = \frac{f(a)^2}{|f'(a)|} = \frac{\exp[-a^2]}{2a}.$$

Let us construct a function

$$g(x) = A(x - u) \exp[-(x - u)^2],$$

that satisfies

$$g(a) = f(a), \quad g'(a) = f'(a).$$

Let $t = a - u$. Solving the resulting system gives

$$t = \frac{a + \sqrt{a^2 + 2}}{2},$$

and

$$p_{lower}(a) = \int_a^\infty g(x)dx = \frac{\exp[-a^2]}{2t} = \frac{\exp[-a^2]}{a + \sqrt{a^2 + 2}}$$

Via construction,

$$p_{lower}(a) < \operatorname{erfc}(a) < p_{upper}(a), \quad a > 0$$

As $t > a$, the denominators satisfy $a + \sqrt{a^2 + 2} > 2a$; therefore, $p_{lower}(a) < p_{upper}(a)$.

Both bounds are explicit and extremely stable in log-space:

$$\begin{aligned} \log[p_{upper}(a)] &= -a^2 - \log(2a), \\ \log[p_{lower}(a)] &= -a^2 - \log[a + \sqrt{a^2 + 2}]. \end{aligned}$$

2.3. Asymptotic Expansions Up to a^{-8}

Let $a \rightarrow \infty$. In classical form (Mills' ratio expansion) we get

$$\operatorname{erfc}(a) \sim \frac{e^{-a^2}}{2a} \left(1 - \frac{1}{2a^2} + \frac{3}{4a^4} - \frac{15}{8a^6} + \frac{105}{16a^8} + O(a^{-10}) \right).$$

Our two bounding expressions admit clean asymptotics in the same form.

The upper bound is in the following form:

$$p_{upper}(a) = \frac{e^{-a^2}}{2a}.$$

This is exactly the leading term of the Gaussian tail asymptotics.

Using

$$p_{lower}(a) = \frac{e^{-a^2}}{a + \sqrt{a^2 + 2}} = \frac{e^{-a^2}}{2a} \cdot \frac{2}{1 + \sqrt{1 + \frac{2}{a^2}}},$$

$\sqrt{1 + \varepsilon} = 1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + \frac{\varepsilon^3}{16} - \frac{5\varepsilon^4}{128} + O(\varepsilon^5)$ is expanded with $\varepsilon = 2/a^2$ and inverted to obtain

$$\frac{2}{1 + \sqrt{1 + \frac{2}{a^2}}} = 1 - \frac{1}{2a^2} + \frac{1}{2a^4} - \frac{5}{8a^6} + \frac{7}{8a^8} + O(a^{-10}).$$

Therefore,

$$p_{lower}(a) \sim \frac{e^{-a^2}}{2a} \left(1 - \frac{1}{2a^2} + \frac{3}{4a^4} - \frac{15}{8a^6} + \frac{7}{16a^8} + O(a^{-10}) \right).$$

These formulations show that both bounds agree with the exact tail of the first two terms: the lower bound differs starting at a^{-4} , whereas the upper bound is the leading term only. In practice, either bound can be used as the starting point for controlled asymptotic refinement.

Table 1 compares $\log[\operatorname{erfc}(a)]$, $\log[p_{lower}(a)]$, and $\log[p_{upper}(a)]$ for $a = 1, \dots, 20$, with all logarithms being natural. As expected, $\log p_{lower}(a) < \log \operatorname{erfc}(a) < \log p_{upper}(a)$, and the gap decreases rapidly as a increases. Accurate Gaussian tail values were obtained using the `scipy.special.erfc` function from the SciPy library (Python).

Table 1. Comparison of the true value logarithms and the lower and upper bounds.

a	$\log[\operatorname{erfc}(a)]$	$\log[p_{lower}(a)]$	$\log[p_{upper}(a)]$
1	-1.9703877476	-2.0050525387	-1.6931471806
2	-5.4857235023	-5.4927894251	-5.3862943611
3	-10.8411452796	-10.8431850133	-10.7917594692
4	-18.1085605497	-18.1093207643	-18.0794415417
5	-27.3216717832	-27.3220104357	-27.3025850930
6	-38.4983434109	-38.4985148043	-38.4849066498
7	-51.6490133313	-51.6491086758	-51.6390573296
8	-66.7802542084	-66.7803112271	-66.7725887222

9	-83.8964521172	-83.8964882130	-83.8903717579
10	-103.0006712625	-103.0006951848	-102.9957322736
11	-124.0951975621	-124.0951493044	-124.0910424534
12	-147.1815171929	-147.1815081064	-147.1780538303
13	-172.2609917239	-172.2610420738	-172.2580965380
14	-199.3346401127	-199.3347458240	-199.3322045102
15	-228.4032738090	-228.4034122328	-228.4011973817
16	-259.4675444399	-259.4676833305	-259.4657359028
17	-292.5279799500	-292.5280861557	-292.5263605246
18	-327.5850125174	-327.5850585883	-327.5835189385
19	-364.6390001601	-364.6389683326	-364.6375861597
20	-403.6902433592	-403.6901271169	-403.6888794541

3. Practical Use and Numerical Stability

Underflow-safe evaluation. The following are always computed in log-space:

$$\log[p_{\text{upper}}(a)] = -a^2 - \log[2a],$$
$$\log[p_{\text{lower}}(a)] = -a^2 - \log[a + \sqrt{a^2 + 2}].$$

Asymptotic refinement. For very large a , we append the known asymptotic correction terms $\text{erfc}(a)$ to a^{-8} for improved accuracy beyond the single-term upper bound.

Compatibility with the standard normal tail. If the one-sided Gaussian tail

$$Q(z) = \int_z^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

is needed, then convert via $a = z/\sqrt{2}$ and scale accordingly:

$$Q(z) = \frac{1}{2} \text{erfc}\left(\frac{z}{\sqrt{2}}\right).$$

The same bounding strategy applies verbatim after this change of variables.

4. Discussion

The proposed tangent and rectangle bracketing method yields a pair of closed-form expressions, which are

1. Simple (two elementary functions of a),
2. Monotone ($p_{\text{lower}} \leq \text{erfc} \leq p_{\text{upper}}$),
3. Asymptotically sharp (both match the leading $\frac{e^{-a^2}}{2a}$, while the lower bound matches the next term $-\frac{1}{2a^2}$),
4. Numerically robust in log-space for a in the meta-analytic tail.

In workflows in which a strict bound is acceptable (e.g., to certify that a meta-analytic p -value is below a given threshold), the lower bound often suffices. For point estimates, either the midpoint in linear space or the log-mean in log-space provides a “free” estimator with guaranteed bracketing error. If higher accuracy is needed, the classical asymptotic series up to a^{-8} may be added to the leading term; the resulting log can be evaluated via $\log \Sigma$ with log-sum-exp to avoid catastrophic cancellation.

5. Conclusions

Herein, we introduce a self-contained analytic approximation of $\log \operatorname{erfc}(a)$ for large a based on a geometric tangent argument. This method yields closed-form upper and lower bounds with asymptotic expansions up to a^{-8} . This is especially useful in meta-analyses in which large z-scores must be handled accurately and safely. This approach is straightforward to implement and integrate into statistical codebases.

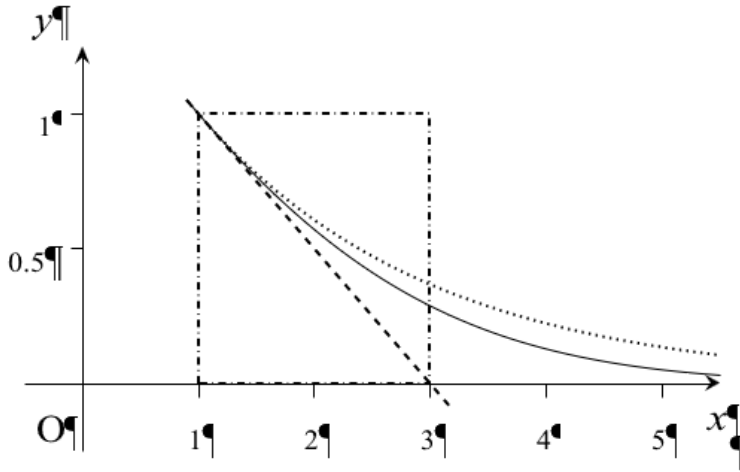


Figure 2. The Gaussian tail (dotted line) is compared with the exponential function and its tangent-based triangular approximation. For the tangency point chosen, the Gaussian tail lies between the exponential decay curve (solid line) and the area of the right triangle generated by the tangent line (dashed line), demonstrating that the Gaussian tail integral is geometrically sandwiched between these two quantities.

Appendix A. Derivation Details

A.1 Solving the Tangent-Matching System

Let $g(x) = A(x - u)e^{-(x-u)^2}$, then

$$g(a) = Ate^{-t^2} = e^{-a^2}, g'(a) = Ae^{-t^2}(1 - 2t^2) = -2ae^{-a^2}, t = a - u.$$

Dividing the second equation by the first yields

$$\frac{1 - 2t^2}{t} = -2a \Leftrightarrow 2t^2 - 2at - 1 = 0 \Leftrightarrow t = \frac{a \pm \sqrt{a^2 + 2}}{2}.$$

We choose the positive root $t = \frac{a + \sqrt{a^2 + 2}}{2} > a$ to ensure g lies below f for $x > a$. Then

$$A = \frac{e^{-a^2}}{te^{-t^2}} = \frac{e^{-(a^2 - t^2)}}{t}, \int_a^\infty g(x) dx = \frac{A}{2} e^{-t^2} = \frac{e^{-a^2}}{2t} = \frac{e^{-a^2}}{a + \sqrt{a^2 + 2}}$$

A.2 Asymptotics of p_{lower} to a^{-8}

Let us write $p_{\text{lower}}(a) = \frac{e^{-a^2}}{2a} \cdot R(a)$ with

$$R(a) = \frac{2}{1 + \sqrt{1 + \frac{2}{a^2}}}$$

Set $\varepsilon = 2/a^2$. Using $\sqrt{1 + \varepsilon} = 1 + \frac{\varepsilon}{2} - \frac{\varepsilon^2}{8} + \frac{\varepsilon^3}{16} - \frac{5\varepsilon^4}{128} + O(\varepsilon^5)$ and expanding $2/(1 + \sqrt{1 + \varepsilon})$ yields

$$R(a) = 1 - \frac{1}{2a^2} + \frac{1}{2a^4} - \frac{5}{8a^6} + \frac{7}{8a^8} + O(a^{-10}),$$

and the claimed series follows.

Acknowledgments: I thank my wife for her suggestions and encouragement. This work was supported by JSPS KAKENHI Grant Numbers JP25K22410 and Yokohama Foundation for Advancement of Medical Science (Medical Digitalization Grant). I thank Editage (www.editage.com) for providing

professional English-language editing services. AI-assisted tools (Microsoft Copilot, 2026) were used solely for idea organization and for improving the clarity and fluency of the English text. All mathematical derivations, theoretical arguments, numerical experiments, and final scientific content were entirely produced and verified by the authors.

Conflicts of Interest: The author declares no competing interests.

References

1. M. M. Shepherd, and J. G. Laframboise, "Chebyshev Approximation of $(1 + 2x) \exp(x^2) \operatorname{erfc} x$ in $0 < x < \infty$," *Mathematics of Computation*, vol. 36, no. 153, pp. 249-253, 1981.
2. W. J. Cody, "Rational Chebyshev Approximations for the Error Function," *Mathematics of Computation*, vol. 23, no. 107, pp. 631-637, 1969.
3. C. Hastings, Jr., *Approximations for Digital Computers*: Princeton University Press, 1955.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.