# Standard Normal Distribution: Efficiency of Approximations

There are several ways to approximate the cumulative distribution function of the standard normal distribution

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

Some of the methods are mentioned in Appendix D2; for Romberg extrapolation see Exercise 1.3. The demands for either speed or accuracy will decide on which code to use. The figure gives an overview.
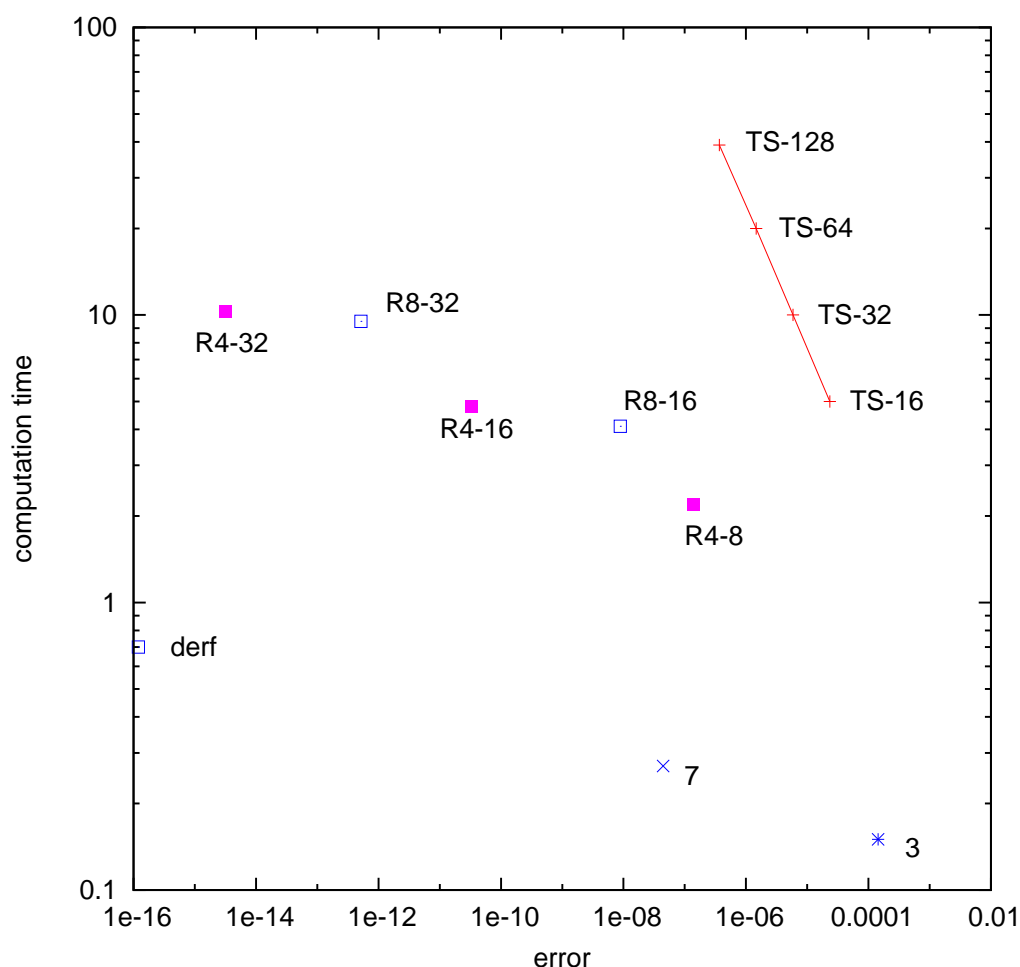
Meaning of notations:

- **derf** generic code calculating the error function in double precision
- **7** analytic approximation with maximum absolute error of 10^{-7}
- **3** analytic approximation with maximum absolute error of 10^{-3}
- **TS-16** trapezoidal sum with 16 subintervals. analogously: TS-32, 64, 128
- **R8-16** Romberg extrapolation, starting with 8 subintervals and terminating with 16. That means one "doubling." analogously: R8-32, R8-64
- **R4-32** Romberg extrapolation, starting with 4 subintervals and terminating with 32. That means three "doublings." analogously: R4-8, R4-16, R4-64

The computation times are taken from a specific notebook, in seconds, but scaled by a factor of 10^8. The error is the average error taken from 10^6 evaluations for 10^6 different values of x. In the figure, the TS-marks of codes belonging to the same family of second-order algorithms are connected by a straight line. Their order $p = 2$ nicely matches the rule

$$\text{costs} \sim (\text{error})^{-1/p}$$

(in the one-dimensional situation); the slope of the straight line is $-1/2$. And the orders of the extrapolations can be seen as well. For example, for the three doublings of R4-16 and R8-32 the marks are on a straight line with slope $-1/6$, which reflects the expected order $p = 6$.

The more efficient codes are in the lower left part, whereas the less efficient codes are in the upper right part of the square.

Conclusions

The double precision code `derf` is both remarkably fast and delivers full accuracy of about 16 decimal digits. Hence, if `derf` is available on a computer, then use it; the other codes are hardly competitive. But in case `derf` is not available, then distinguish between two situations: For moderate accuracy, say up to 7 decimal digits, use the analytic approximation formula indicated by the "7" in the figure (which is even faster than `derf`). In this range of accuracy, Romberg extrapolation is comparably too expensive. But if high accuracy is needed, then a price for a more involved code such as R4-32, for example, must be paid. We note in passing that there are even faster analytic approximations for low accuracy (example "3" in the figure), but probably of little use.

Sources

Analytic approximation formulas can be found in [M. Abramowitz & I.A. Stegun: Handbook of Mathematical Functions], for formula "7" see also Appendix D2. The trapezoidal sum for $F(x)$, with n subintervals, nodes $x_i = ih$ (for $x > 0$, $h := \frac{x}{n}$, $i = 0, 1, \dots, n$) is TS-n,

$$\frac{x}{n} \left(\frac{1}{2} f(0) + f(x_1) + \dots + f(x_{n-1}) + \frac{1}{2} f(x) \right) \quad \text{for } f(t) := \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2}),$$

see also Appendix C1. Extrapolation repeats the calculation for one or more doubled values of n, interpolates the approximations TS-n, TS-$2n$, ... by a polynomial in h^2, and extrapolates to $h = 0$.

For negative arguments use $F(x) = 1 - F(-x)$.

If the intrinsic error function `derf` is used [W.J. Cody: Rational Chebyshev approximation for the error function. Math. Comp. **22** (1969) 631–637], then for $x > 0$

$$F(x) = \frac{1}{2} \left(1 + \text{erf}(\frac{x}{\sqrt{2}}) \right).$$