Areas under the normal curve

In order to use the normal curve to approximate discrete probabilities we have to be able to calculate values of

$$\frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2}\, dt \,.$$

But any table we use can only store values for a finite range of x. I have chosen $x = 5$ somewhat arbitrarily. What to do for $x > 5$?

There is a useful trick. First of all, we know that

$$\frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-t^2/2}\, dt = 1/2$$

so

$$\frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2}\, dt = (1/2) - \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2}\, dt \,.$$

But integration by parts tells us that

$$\int u\,dv = uv - \int v\,du$$

and in our case

$$\int_x^\infty e^{-t^2/2}\, dt = \left[\frac{e^{-t^2/2}}{t} \right]_x^\infty - \int_x^\infty \frac{e^{-t^2/2}}{t^2}\, dt$$

$$= \frac{e^{-x^2/2} /}{x} - \int_x^\infty \frac{e^{-t^2/2}}{t^2}\, dt$$

$$< \frac{e^{-x^2/2}}{x} \,.$$

But also

$$\int_x^\infty \frac{e^{-t^2/2}}{t^2}\, dt = \left[\frac{e^{-t^2/2}}{t^3} \right]_x^\infty - \int_x^\infty e^{-t^2/2} \frac{3}{t^4}\, dt$$

$$= \frac{e^{-x^2/2}}{x^3} - \int_x^\infty e^{-t^2/2} \frac{3}{t^4}\, dt$$

$$< \frac{e^{-x^2/2}}{x^3} \,.$$

Therefore

$$\int_x^\infty e^{-t^2/2}\, dt > \frac{e^{-x^2/2}}{x} - \frac{e^{-x^2/2}}{x^3} = e^{-x^2/2} \left(\frac{1}{x} - \frac{1}{x^3} \right) \,.$$

To summarize, we can use this inequality as an estimate:

$$\frac{e^{-x^2/2}}{\sqrt{2\pi}} \left(\frac{1}{x} - \frac{1}{x^3} \right) < \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2}\, dt < \frac{e^{-x^2/2}}{\sqrt{2\pi}} \left(\frac{1}{x} \right) \,.$$

1

For example when $x = 5$:

$$0.0000002855 < \frac{1}{\sqrt{2\pi}} \int_5^\infty e^{-t^2/2}\, dt < 0.0000002973 \,.$$

This process can be continued, and leads to the fact that the integral lies between any two successive partial sums of the series

$$e^{-x^2/2} \left(\frac{1}{x} - \frac{1}{x^3} + \frac{3}{x^5} - \frac{3 \cdot 5}{x^7} + \frac{3 \cdot 5 \cdot 7}{x^9} - \cdots \right) \,.$$

This series is rather strange. It does not converge, but the first few terms will decrease in magnitude, and it may therefore be used for computation.

When $x = 4$, for example, we get for the series

n	t_n	s_n	$s_n e^{-x^2/2}$
0	0.25000	0.2500	0.00008387
1	-0.01563	0.2344	0.00007862
2	0.00293	0.2373	0.00007961
3	-0.00092	0.2364	0.00007930
4	0.00040	0.2368	0.00007943
5	-0.00023	0.2366	0.00007936
6	0.00015	0.2367	0.00007941
7	-0.00013	0.2366	0.00007937
8	0.00012	0.2367	0.00007941
9	-0.00013	0.2366	0.00007937
10	0.00015	0.2367	0.00007942

The series starts to diverge here, and the most accurate possible value you can get in this way is 'is between 0.00007937 and 0.00007941.