Linear Interpolation

Linear interpolation is a method of approximating the graph of a function in between two points on the graph by the straight line between them. Provided the graph does not deviate too far from the line between the two points, we can use the coordinates of the points on the line as approximations of the coordinates of the points on the graph directly above or below them. This is useful when we do not have a formula for the function, as in the case of the normal cdf. Linear interpolation can be used to approximate "in between" values in tables such as the normal tables. In the days before modern hand-held calculators, people commonly used linear interpolation to get such "in between" values in trig, logarithmic, and statistical tables; a value obtained in this manner is usually better than the closest value in the table.

When using the normal cdf tables, we sometimes wish to interpolate to obtain a probability given a value of z which lies between two consecutive values of z in the margins of the table, and other times we wish to interpolate to find the value of z which corresponds to a probability lying between two adjacent probabilities in the body of the table. Using the normal tables in the text, we'll look at one example of each type to illustrate the process. In both examples, Z represents a standard normal random variable. The text's normal tables give the values of the standard normal distribution's ***survival function***: for any random variable Y with cdf $F(y)$, its survival function is the function $S(y) = 1 - F(y)$.

Example 1. Use linear interpolation to find $P(Z > 1.263) = S(1.263)$.

Since 1.263 lies between the adjacent values 1.26 and 1.27 in the table margins, consider the line segment on the graph of the survival function of Z between the two points $(1.26, S(1.26) = 0.1038)$ and $(1.27, S(1.27) = 0.1020)$. The point on this line with first coordinate 1.263 will be the one directly above or below the point $(1.263, S(1.263))$ on the graph of the survival function, and we will use the second coordinate of that point (call it p) to approximate $P(Z > 1.263) = S(1.263)$.

So consider the three points $(1.26, 0.1038)$, $(1.263, p)$, and $(1.27, 0.1020)$ on this line. Since a straight line has constant slope, if we calculate the slope between any two of these points we will get the same result as we do if we use another pair. Thus

$$\frac{p - 0.1038}{1.263 - 1.26} = \frac{0.1020 - 0.1038}{1.27 - 1.26}$$

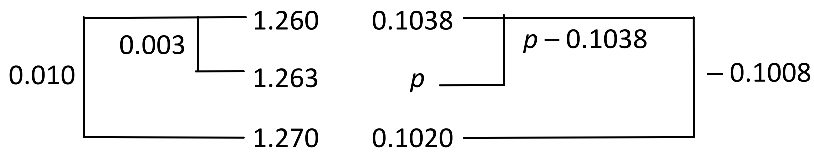
$$\frac{p - 0.1038}{0.003} = \frac{0.1020 - 0.1038}{0.01}$$

$$p - 0.1038 = \frac{0.003}{0.010}(0.1020 - 0.1038)$$

$$p = 0.1038 + \frac{3}{10}(0.1020 - 0.1038) = 0.1038 + \frac{3}{10}(-0.0018)$$

I like to think of this as saying that since 1.263 is 3/10 of the way from 1.26 to 1.27, $P(Z > 1.263)$ will lie approximately 3/10 of the way from $P(Z > 1.26) = 0.1038$ to $P(Z > 1.27) = 0.1020$.

The previous discussion shows the reason for the calculations used in linear interpolation. When actually doing it, I like to set it up as follows:



So, $p = 0.1038 + (3/10)\,(-0.0018) = 0.1038 - 0.00054 = 0.1033.$

Example 2. Use linear interpolation to find the value c such that $P(Z > c) = 0.3961$.

The probability value 0.3961 lies between $0.3974 = P(Z > 0.26)$ and $0.3936 = P(Z > 0.27)$ in the text's normal table.

So, $c = 0.260 + (13/38)\,(0.010) = 0.260 + 0.0034 = 0.263.$