NOTE ON THE NORMAL PROBABILITY CURVE

BY

BUCKNER SPEED*

Technical Expert, Patent Department, Bell Telephone Laboratories

In a former paper ("A New Graphic Method of Using the Normal Probability Curve" in the *Proceedings* of this Society, Volume II, page 120) I called attention to the fact that if the equation of the normal probability curve is written

$$y = \frac{n}{s\sqrt{2\pi}} e^{-(x^2/2s^2)}$$

and we take logarithms of both sides we get

$$\log y = \log \frac{n}{s\sqrt{2\pi}} - \frac{x^2}{2s^2}$$

and putting $\log y = u$ and $\log \frac{n}{s\sqrt{2\pi}} = k$

$$u = k - \frac{x^2}{2s^2}$$

which is a parabola with its vertex at $\left\{ \begin{array}{l} u = k \\ x = 0 \end{array} \right.$

Now if A, B, C, D, E, F, etc., are equidistant ordinates of the normal distribution curve, and a, b, c, d, e, f, etc., are the common logarithms of the above numbers, then

$$(a - b) - (b - c) = (b - c) - (c - d)$$

which is equivalent to saying that the second order of differences is a constant *i. e.*, a parabola
or

$$d = a - 3b + 3c$$

*Paper presented by Buckner Speed, Technical Expert, Patent Department, Bell Telephone Laboratories, on invitation of the Committee on Program.

or going back to natural numbers

$$\text{2nd term,} \quad D = \frac{A\ C^3}{B^3}$$

and by similar substitutions

$$\text{3rd term,} \quad E = \frac{A^3\ C^6}{B^8}$$

and if in the series A, B, C, D, E, F, etc., we call B the 0 term and C the 1st term, we get

$$n\text{th term,} = \frac{A^{\frac{1}{2}(n^2 - n)}\ C^{\frac{1}{2}(n^2 + n)}}{B^{(n^2 - 1)}}$$

Thus, any portion of the normal distribution curve being known, so that we may measure or otherwise determine three equidistant ordinates, we may extrapolate in either direction.

Now, if we choose A, B, and C, so that B is the maximum ordinate, that is to say, the vertical axis of the curve, and A and C are equidistant from B, and therefore equal to each other, then if we make $\frac{A}{B} = r$, we get

$$D = A\ r^3$$
$$E = A\ r^8$$
$$n^{th} = A\ r^{(n^2 - 1)}$$

The above equations are based on the fact that the second order of differences of a, b, c, d, e, f, g is a constant or that the third order of differences is zero, which is merely another way of saying that the points a, b, c, d, e, f, g are on a parabola.

Now as is well known many collections of data when platted form a skew distribution in which the maximum ordinate is not in the center, and the slopes of the two shoulders is different.

Let us take the case where the third order of differences of a, b, c, d, e, f, g, is constant, and the fourth order is zero.

We then get

$$[(a-b)-(b-c)]-[(b-c)-(c-d)]=[(b-c)-(c-d)]-[(c-d)-(d-c)]$$

from which $e = 4\,b + 4\,d - a - 6\,c$

or going back to natural numbers

$$\text{2nd term,} \quad E = \frac{B^4 D^4}{A\ C^6}$$

$$\text{3rd term,} \quad F = \frac{B^{15} D^{10}}{A^4\ C^{20}}$$

$$\text{4th term,} \quad G = \frac{B^{36} D^{20}}{A^{10}\ C^{45}}$$

$$n\text{th term} = \frac{B^{\frac{1}{2}n(n-1)(n+2)}\ D^{\frac{1}{6}n(n+1)(n+2)}}{A^{\frac{1}{6}n(n^2-1)}\ C^{\frac{1}{2}(n^2-1)(n+2)}}$$

or if we put $\dfrac{B^3 D}{A\ C^3} = M$

$$\text{2nd term,} \quad E = M\ \frac{B\ D^3}{C^3}$$

$$\text{3rd term,} \quad F = M^4\ \frac{B^3 D^6}{C^8}$$

$$\text{4th term,} \quad G = M^{10}\ \frac{B^6\ D^{10}}{C^{15}}$$

$$n\text{th term} = M^{1/6(n^3-n)}\ \frac{B^{1/2(n^2-n)}\ D^{1/2(n^2+n)}}{C^{(n^2-1)}}$$

the coefficient M has the significance of a modulus of skewness, thus $\log\left(\dfrac{1}{M}\right) = k$ where k is the third order of differences of a, b, c, d, e, f, g, that is

$$k = [(a - b) - (b - c)] - [(b - c) - (c - d)]$$

When the probability is normal, and the logarithm curve is a parabola $k = 0$ (*i. e.*, 3rd order of difference) and hence $M = 1$.