Also

$$\frac{1}{\sqrt{2\pi}}\int_0^k e^{-\frac{1}{2}y^2}\,dy = \frac{(a+b)-(c+d)}{2N} = .187, \text{ and } k = .4874.$$

Then,

$$H = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}h^2} = .3635, \text{ and } K = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}k^2} = .3543.$$

All the quantities except r in the following approximate equation are known:

$$\frac{ad-bc}{N^2HK} = r + \frac{r^2}{2}hk + \frac{r^3}{6}(h^2-1)(k^2-1) + \frac{r^4}{24}h(h^2-3)k(k^2-3) + \frac{r^5}{125}(h^4-6h^2+3)(k^4-6k^2+3).$$

Therefore,

$$.0261r^5 + .0681r^4 + .1034r^3 + .1052r^2 + r - .4314 = 0.$$

Then, r is approximately equal to .4051. Consequently, for practical purposes we can assume that $r = .4$.

28 Boody Street
Brunswick, Maine

John L. Roberts

NOTE ON THE DERIVATION OF THE MULTIPLE CORRELATION COEFFICIENT

Consider N observed ~~values~~ of each of n variables. These $n \cdot N$ values may be tabulated in a double-entry table as follows:

$$\begin{array}{c} X_{11}\ X_{12}\ X_{13}\ \cdots\ X_{1N} \\ X_{21}\ X_{22}\ X_{23}\ \cdots\ X_{2N} \\ \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \\ X_{n1}\ X_{n2}\ X_{n3}\ \cdots\ X_{nN} \end{array}$$

where X_{ik} is the k^{th} value of the i^{th} variable.

Using the i^{th} variable as the dependent variable, the general linear relationship between the n variables may be expressed by

$$x_i = {}_ia_1\,x_1 + {}_ia_2\,x_2 + \cdots + {}_ia_{i-1}\,x_{i-1} + {}_ia_{i+1}\,x_{i+1} + \cdots + {}_ia_n\,x_n \qquad (1)$$

where

${}_ia_j$ is the general parameter which is to be determined empirically;

$x_j = X_j - M_j$;

M_j is the arithmetic mean of the j^{th} variable.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org