PHI-COEFFICIENT

The phi-coefficient is actually a product—moment coefficient of correlation* and is a variation of Pearson's definition of r when the two states of each variable are given values of 0 and 1 respectively.

The phi-coefficient was designed for the comparison of truly dichotomous distributions, i.e., distributions that have only two points on their scale which indicate some unmeasurable attribute. Attributes such as living or dead, black or white, accept or reject, and success or failure are examples. It is also sometimes known as the Yule ϕ [1].

If certain allowances are made for continuity, the technique may be applied to observations grouped into two arbitrary, but clearly defined, divisions. It is clear that in many such divisions point attributes are achieved by a binary decision process employing demarcation lines through "grey" or "ill-defined" regions.

The phi-coefficient relates to the 2×2 table*:

	Attribute 1	
Attribute 2	Yes	No
Yes	a	b
No	c	d

If a, b, c, and d represent the frequencies of observation, then ϕ is determined by the relationship

$$\phi = \frac{ad - bc}{\sqrt{\{(a+b)(c+d)(a+c)(b+d)\}}}$$

The phi-coefficient is particularly used in psychological and educational testing where the imposing of a dichotomy on a continuous variable is a frequent occurrence; variables where pass/fail categories are obtained in relation to a threshold score are typical (*see* PSYCHOLOGICAL TESTING THEORY).

It bears a relationship to χ^2, where

$$\phi^2 = \frac{\chi^2}{N} \quad \text{or} \quad \chi^2 = N\phi^2$$

and $N = a + b + c + d$ (*see* CHI-SQUARE TEST—I).

The significance of ϕ may be tested by determining the value of χ^2 from the above relationship and testing in the usual way.

As an example, 43 persons were asked if they believed that there was any truth in horoscopes or in the existence of UFOs. The results gave

	Horoscopes	
UFO's	Some truth	No truth
Might exist	14	10
Don't exist	6	13

Applying the above formula, $\phi = 0.266$.

This value of ϕ corresponds to a value of $\chi^2$of $43 \times (0.266)^2 = 3.04$. This may then be tested against the relevant value of χ^2 for 1 degree of freedom.

An alternative significance test (rarely used) may be performed by considering the standard error of ϕ. Calculation of this is laborious but if N is not too small, then $1/\sqrt{N}$ approximates to it [2].

REFERENCES

1. Yule, G. U. (1912). *J. R. Statist. Soc.*, **75**, 576–642. (On the methods of measuring the association between two variables. The first identification of the phi-coefficient.)
2. McNemar, Q. (1962). *Psychological Statistics*. Wiley, New York. (Justification after use of approximation.)

See also BISERIAL CORRELATION; CORRELATION; PHI-MAX COEFFICIENT; and TWO-BY-TWO (2×2) TABLES.

O. B. CHEDZOY