Biometrika Trust

On the Probable Error of a Partial Correlation Coefficient

Author(s): David Heron

Source: *Biometrika*, Vol. 7, No. 3 (Apr., 1910), pp. 411-412

Published by: [Biometrika Trust](#)

Stable URL: <http://www.jstor.org/stable/2345396>

Accessed: 18/06/2014 20:17

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Biometrika Trust is collaborating with JSTOR to digitize, preserve and extend access to *Biometrika*.

<http://www.jstor.org>

BIOMETRIC NOTES.

I. An Abac for determining the Probable Errors of Correlation Coefficients.

By DAVID HERON, M.A.

THE probable errors of correlation coefficients are so often required that the accompanying abac which enables them to be determined at once will save considerable labour. From this abac the probable errors can be read off correct to at least two decimal places; greater accuracy is seldom required.

The principle on which the abac is constructed is simple. If we take the well-known formula for the probable error of a correlation coefficient

$$E_r = \frac{.67449 (1 - r^2)}{\sqrt{n}},$$

where r is the correlation and n the number of observations, and express it in logarithmic form, we get

$$\log (E_r) = \bar{1} \cdot 8289755 + \log (1 - r^2) - \frac{1}{2} \log n.$$

For any constant value of r there is thus a linear relationship between the logarithms of E_r and n . Hence by plotting E_r and n on logarithmic scales and ruling lines for a sufficient number of values of r , we can find E_r from r and n without difficulty.

The use of the abac can best be illustrated by an example. Let the correlation between two variables be .55 and the number of observations 160; then to find the probable error of r , we must read along the perpendicular line from the number 160 on the scale of frequency until it crosses the diagonal line representing a correlation of .55. The position of the point of intersection of these two lines is then read, by aid of the vertical lines, on the scale of probable errors and the value .037 so obtained.

After the diagram had been drawn in pencil, the whole of the laborious work of ruling in the lines in ink and preparing the diagram for reproduction was undertaken by Miss H. G. Jones, and I have to thank her most heartily for her careful work.

II. On the Probable Error of a Partial Correlation Coefficient.

By DAVID HERON, M.A.

In a paper "On the Theory of Correlation for any Number of Variables, treated by a New System of Notation,"* G. Udny Yule has suggested that the probable error of a partial correlation coefficient, giving the correlation between two variables for a constant value of a third, is of the same form as the probable error of the direct correlation between any two variables, but his proof by reason of its very general nature is by no means easy to follow and it seems desirable

* *Proc. R.S.*, Vol. 79, pp. 182 et seq.

to investigate the probable error of this coefficient on the lines laid down by Pearson and Filon in their classic paper on the probable errors of frequency constants*.

Let ${}_{3\rho_{12}}$ be the correlation between variables 1 and 2 for a constant value of 3.

Then

$${}_{3\rho_{12}} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}}.$$

Taking logarithmic differentials, we find,

$$\frac{d{}_{3\rho_{12}}}{{}_{3\rho_{12}}} = \frac{dr_{12}}{r_{12} - r_{13}r_{23}} + \frac{dr_{13}(r_{13}r_{12} - r_{23})}{(r_{12} - r_{13}r_{23})(1-r_{13}^2)} + \frac{dr_{23}(r_{23}r_{12} - r_{13})}{(r_{12} - r_{13}r_{23})(1-r_{23}^2)}.$$

Squaring, summing and dividing both sides by n , the total number of observations, and substituting for the standard deviation of the errors in r_{12} , r_{13} , r_{23} , and for the correlations of errors in those quantities the values given by Pearson and Filon in the paper already cited, we find after striking out the common factors $(1-r_{13}^2)(1-r_{23}^2)$

$$\begin{aligned} & \frac{\Sigma {}_{3\rho_{12}}^2 (r_{12} - r_{13}r_{23})^2 (1-r_{13}^2)(1-r_{23}^2)}{{}_{3\rho_{12}}^2} \cdot n \\ &= (1-r_{12}^2-r_{13}^2-r_{23}^2+2r_{12}r_{13}r_{23})\{(1-r_{23}^2)(r_{23}-r_{12}r_{13})r_{12}r_{13} \\ & \quad + (1-r_{13}^2)(r_{13}-r_{12}r_{23})r_{12}r_{23} - (r_{13}-r_{12}r_{23})(r_{23}-r_{12}r_{13})r_{13}r_{23}\} \\ & \quad + (1-r_{13}^2)(1-r_{23}^2)\{(1-r_{12}^2)+(r_{13}-r_{12}r_{23})^2+(r_{23}-r_{12}r_{13})^2-2r_{23}(1-r_{12}^2)(r_{23}-r_{12}r_{13}) \\ & \quad - 2(1-r_{12}^2)(r_{13}-r_{12}r_{23})r_{13}+2(r_{13}-r_{12}r_{23})(r_{23}-r_{12}r_{13})r_{12}\}. \end{aligned}$$

But the second part of the expression to the right of the sign of equality reduces to

$$(1-r_{13}^2)(1-r_{23}^2)(1-r_{12}^2)(1-r_{12}^2-r_{13}^2-r_{23}^2+2r_{12}r_{13}r_{23}),$$

and the second factor of the first part reduces to

$$1-r_{12}^2-r_{13}^2-r_{23}^2+2r_{12}r_{13}r_{23}-(1-r_{12}^2)(1-r_{13}^2)(1-r_{23}^2),$$

so that the whole expression on the right reduces at once to

$$(1-r_{12}^2-r_{13}^2-r_{23}^2+2r_{12}r_{13}r_{23})^2.$$

But

$${}_{3\rho_{12}}^2 = \frac{(r_{12}-r_{13}r_{23})^2}{(1-r_{13}^2)(1-r_{23}^2)},$$

so that

$$\begin{aligned} \Sigma {}_{3\rho_{12}}^2 &= \frac{1-r_{12}^2-r_{13}^2-r_{23}^2+2r_{12}r_{13}r_{23}}{\sqrt{n}(1-r_{13}^2)(1-r_{23}^2)} \\ &= \frac{1-{}_{3\rho_{12}}^2}{\sqrt{n}}, \end{aligned}$$

and the probable error of ${}_{3\rho_{12}}$ is

$$= \frac{.67449(1-{}_{3\rho_{12}}^2)}{\sqrt{n}},$$

and is thus, as stated by Yule, of the same form as the probable error of the direct correlation between any two variables.

* "On the Probable Errors of Frequency Constants and on the Influence of Random Selection on Variation and Correlation." By Karl Pearson, F.R.S., and L. N. G. Filon, B.A., *Phil. Trans. R.S., A*, Vol. 191, pp. 229 et seq.