Note on Estimating the Relative Influence of Two Variables upon a Third
Author(s): R. H. Hooker and G. U. Yule
Source: *Journal of the Royal Statistical Society*, Vol. 69, No. 1 (Mar., 1906), pp. 197-200
Published by: Wiley for the Royal Statistical Society
Stable URL: https://www.jstor.org/stable/2339552
Accessed: 30-10-2019 17:43 UTC

MISCELLANEA.

CONTENTS:

I.—*Note on Estimating the Relative Influence of Two Variables upon a Third.* By R. H. Hooker and G. U. Yule.

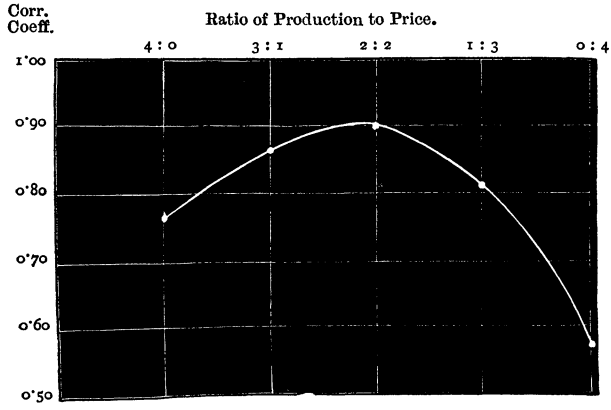
The correlation coefficient usually furnishes a simple test of whether one phenomenon depends upon another. Generally, however, any particular phenomenon in economics is due not to a single cause, but to a multiplicity of causes, of which it is not always easy to say which is the chief. Consider for instance the following table, showing for the past fifteen seasons the exports of wheat from India, together with the production of that cereal there, and the Gazette price of wheat in England. Inspection of this table shows that a large exportation of wheat does not invariably follow a large production, but that in the latter case the price turns out to be low. Clearly therefore price as well as production is an important factor in the wheat export trade of India—more so than might perhaps have been expected—and it is accordingly an interesting problem to determine their relative importance.

Table showing the Production and Exportation of Wheat in India, and Gazette Price of Wheat in England.

Years ending 31st March,	Production.	Exports of Subsequent Twelve Months.	Gazette Price of Wheat.*	
	Tons.	Tons.	*s.*	*d.*
1889-90	6,123,000	716,025	31	11
'90-91	6,876,000	1,515,335	37	–
'91-92	5,535,000	748,673	30	3
'92-93	7,193,000	607,843	26	4
'93-94	6,771,000	344,506	22	10
'94-95	6,279,334	500,209	23	1
'95-96	5,510,965	95,531	26	2
'96-97	4,892,879	119,630	30	2
'97-98	6,686,778	976,025	34	–
'98-99	6,339,603	485,204	25	8
'99-1900	4,869,727	2,501	26	11
1900-01	6,765,717	366,091	26	9
'01-02	6,063,506	514,607	28	1
'02-03	7,766,096	1,295,566	26	9
'03-04	9,601,393	2,150,025	28	4

* Calendar years 1890-1904.

I. An elementary method which suggested itself as offering a solution was the following: The quantities produced and exported and the prices were reduced to index numbers, and the exports were then correlated (i) with the production, (ii) with the price. The index numbers of production and price were then added together, and (iii) the resultant figure correlated with exports. Finally, the exports were correlated with index numbers weighted by taking production and price in the proportion of (iv) 3 to 1 and (v) 1 to 3. Five coefficients of correlation were thus obtained between exports and figures which may be regarded as representing production and price in the ratios 4 : 0, 3 : 1, 2 : 2, 1 : 3, 0 : 4. The coefficients thus obtained* were + 0·77, 0·86, 0·90, 0·81 and 0·58 respectively. Plotting these on a diagram, and drawing a smooth curve through the five points, this curve shows a maximum just to the left of the central ordinate.



It seems a fair inference that this maximum indicates the relative influence of the two factors. It may therefore be concluded that during the past fifteen years production has been of slightly greater importance than price in determining the changes in the quantity of wheat exported from India from year to year. The influence of the price that Great Britain is willing to pay seems much greater than might have been anticipated, and argues unusual capability on the part of India to substitute some other food for wheat if only the price be satisfactory.

II. The foregoing method is, however, clumsy and becomes somewhat lengthy if the observations are at all numerous; and the general theory of correlation leads to the same result in a

* The correlation coefficients, not of the usual differences from the mean but of the successive annual differences, were calculated, as shown in the last number of this *Journal*. (Hooker: "On the Correlation of Sequences of Observations," vol. lxviii, December, 1905, p. 697.) The fall in prices during the period would somewhat mask the connection between price and export were the differences from the mean to be correlated.

simpler way. For the preceding method is equivalent to making the correlation coefficient between x_1 and $x_2 + bx_3$ a maximum by varying b.

Let $$x_2 + bx_3 = z,$$

then, adopting the usual notation—*

$$\Sigma(x_1 z) = \Sigma(x_1 x_2) + b\Sigma(x_1 x_3) = n(r_{12}\sigma_1\sigma_2 + br_{13}\sigma_1\sigma_3)$$

and $$\Sigma z^2 = n(\sigma_2^2 + b^2\sigma_3^2 + 2br_{23}\sigma_2\sigma_3).$$

Hence the correlation coefficient

$$r_{x_1 z} = \frac{r_{12}\sigma_2 + br_{13}\sigma_3}{\sqrt{\sigma_2^2 + b^2\sigma_3^2 + 2br_{23}\sigma_2\sigma_3}}$$

To find the value of b for which this is a maximum, differentiate with regard to b, and equate to zero; then

$$\frac{r_{13}\sigma_3}{\sqrt{\sigma_2^2 + b^2\sigma_3^2 + 2br_{23}\sigma_2\sigma_3}} - \frac{(r_{12}\sigma_2 + br_{13}\sigma_3)(b\sigma_3^2 + r_{23}\sigma_2\sigma_3)}{(\sigma_2^2 + b^2\sigma_3^2 + 2br_{23}\sigma_2\sigma_3)^{\frac{3}{2}}} = 0;$$

$$r_{13}\sigma_3\sigma_2^2 + br_{13}r_{23}\sigma_2\sigma_3^2 - br_{12}\sigma_2\sigma_3^2 - r_{12}r_{23}\sigma_2^2\sigma_3 = 0;$$

whence $$b = \frac{(r_{13} - r_{12}r_{23})\sigma_2}{(r_{12} - r_{13}r_{23})\sigma_3}.$$

Hence the relative influence of the two factors may be found by calculating the three correlation coefficients between the three variables and then finding the value of $\frac{(r_{13} - r_{12}r_{23})\sigma_2}{(r_{12} - r_{13}r_{23})\sigma_3}$. That is to say, the relative influences of x_2 and x_3 on x_1, as determined by method I, are proportionate to the "net" or "partial" regressions"† of x_1 on x_2 and of x_1 on x_3, viz.:—

$$\frac{r_{12} - r_{13}r_{23}}{1 - r_{23}^2} \cdot \frac{\sigma_1}{\sigma_2} \text{ and } \frac{r_{13} - r_{12}r_{23}}{1 - r_{23}^2} \frac{\sigma_1}{\sigma_3}.$$

Method I thus furnishes an interesting, if somewhat lengthy, means of determining graphically the ratio of the net regressions.

In the particular case under consideration this works out to $\frac{\text{price}}{\text{production}} = 0{\cdot}97$; indicating, as already found, very nearly equality of influence between price and production.

Upon comparison with the article last referred to, it will be seen that the maximum value of $r_{x_1 z}$ is identical with the quantity which Mr. Yule denotes by the symbol R,‡ and for which he suggests the name "coefficient of double correlation." It is given at once by inserting the value of b, as determined above, in the expression for $r_{x_1 z}$. In the present case we have

$$\mathrm{R} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{23}r_{31}}{1 - r_{23}^2}} = +0{\cdot}905.$$

* Yule: "On the Theory of Correlation, case of Three Variables," *Journal*, vol. lx, December, 1897, pp. 831—2.

† *Ibid.*, p. 832.

‡ *Ibid.*, p. 833.

No practical use has, we believe, been made of this coefficient of double correlation, but the above reasoning appears to indicate that it is of considerable importance, and that it may be regarded as indicative of the closeness of the causal connection between one variable and the joint influence of two other variables of which the first is a function.

II.—*An Analysis of Australian Imports of Merchandise from the United Kingdom and Foreign Countries.* By THOMAS H. HAYNES.

THE course of British trade with the Colonies of late years is attracting considerable interest, and has been the subject of much discussion. Some difficulty arises in tracing the facts, and thus in appreciating the actual tendencies of this trade owing to the method on which statistical abstracts are compiled. Bullion and specie are often included with merchandise, and no attempt, as a rule, is made to distinguish between goods in regard to which there is foreign competition and others.

In the case of Canada no analysis of imports can be made, raw material and manufactures thereof not being separated, but a glance at the totals is instructive.

TABLE 1.

[00,000's omitted.]

	From United Kingdom.		Total Imports of Merchandise.		From United Kingdom.		Total Imports of Merchandise.
	$	Per cent.	$		$	Per cent.	$
1888.....	39,1	36·0	108,7	1897.....	29,3	25·6	114,5
'89.....	42,2	36·8	114,6	'98.....	32,4	23·9	135,9
'90.....	43,3	35·8	120,8	'99.....	36,8	23·3	158,0
'91.....	42,0	35·6	118,1	1900.....	44,9	24·8	181,3
'92.....	41,0	32·7	125,6	'01.....	42,9	23·0	186,8
'93.....	42,6	34·8	122,5	'02.....	42,2	23·9	295,9
'94.....	37,0	31·0	119,4	'03	58,9	25·3	232,2
'95.....	31,0	29·2	106,2	'04.....	61,9	24·6	251,3
'96.....	33,0	29·2	112,7				

Preferential treatment was accorded to the Mother Country in 1898, the year which marks the commencement of increasing trade and an arrest of the decline in imports from the United Kingdom. If the figures for imports entered for consumption only are taken, the British percentage comes out slightly higher, but the relative position remains the same. A recovery in amount has been made since 1898, but not much in percentage.

As regards Australia, the statistical abstract issued in London gives the Commonwealth's imports of merchandise and bullion all