*Statistical Science*
1995, Vol. 10, No. 4, 364–376

# Correlations Genuine and Spurious in Pearson and Yule

**John Aldrich**

*Abstract.* Karl Pearson and G. Udny Yule developed the main interpretations of correlation used by statisticians for the past century or so. They also examined a number of situations in which the correlation inference was unsatisfactory. This paper considers the development of their ideas on both genuine and spurious correlations and makes some reference to related modern work.

*Key words and phrases:* Pearson, Yule, Student, correlation, causation, spurious correlation, spurious regression, Simpson's paradox, variate-difference correlation method.

## 1. INTRODUCTION

Francis Galton invented correlation, but Karl Pearson was chiefly responsible for its development and promotion as a scientific concept of universal significance. This paper treats the development of his ideas and those of his one-time assistant G. Udny Yule. Between them they made correlation analysis.

The emergence of correlation was one of the main developments in statistics during the late 19th century. Galton (1890) gave his own account and the subject has been well served by historians—including MacKenzie (1981), Porter (1986) and Stigler (1986). There are biographical studies of Pearson and Yule by E. S. Pearson (1936, 1938), Eisenhart (1974) and Yates (1952).

The present account emphasises the way Pearson and Yule interpreted correlations. Pearson's understanding of the relationship between correlation and causation is discussed in Section 2. The first paradigm of correlation, as an expression of dependence on common causes, is treated in Section 3. The second, as an expression of direct causal relationship—Yule's contribution—is taken up in Section 5. Section 8 considers how these paradigms were transferred to the time series context.

The other sections treat varieties of "spurious correlation," which receive so much attention because there are more ways of going wrong than going right and because each case provoked reflection on the nature of "genuineness" or what it is that correlation signifies in nonpathological cases. Pearson discovered—or invented—spurious correlation very soon after starting work on correlation and he kept finding fresh instances. Elderton's (1907, page 122) Pearsonian textbook described the problem: "it is possible to obtain a significant value for a coefficient of correlation when in reality the two functions are absolutely uncorrelated." In the modern literature "spurious correlation" is applied to several distinct cases, including some first discussed by Yule. Actually he avoided the term and used "spurious" only when quoting Pearson.

Sections 3 and 4 treat "spurious correlation due to the use of ratios." This, the first case of spurious correlation, has never disappeared from the literature: Kronmal (1993) reviewed recent contributions to the regression model variant. Spurious correlation through "mixing of races" (Section 6) is an ancestor of "Simpson's paradox." This paradox is well known among statisticians (see, e.g., Wagner, 1982), but it has also been discussed by philosophers of science (e.g., Cartwright, 1983). Yule's "illusory" association (Section 7) is related to the "spurious correlation" of Simon (1954): the correlation does not indicate a direct causal relation, but common dependence on a third variable. Yule's "nonsense-correlations" of time series (Section 11) are related to Granger and Newbold's (1974) "spurious regression." Here correlation is taken to indicate dependence when there is none—not even indirect dependence. The same cases reappear in discussions of the general issue of causal inference (see, e.g., Spirtes, Glymour and Schienes, 1993).

*John Aldrich is a Lecturer in the Department of Economics, University of Southampton, England.*

Of course Pearson and Yule were not the first to study the problem of interpreting associations: related issues were treated in the literature on scientific method and probability. However, Edgeworth was the only correlation pioneer who paid much attention to this literature; see Section 11.

The account that follows is of big disagreements between Pearson and Yule. It does not cover all their differences. The one over measures of association, which they saw as deeper than any treated here, is only mentioned. In his memoir of Pearson, Yule (1936, page 84) recalled that dispute, which consumed hundreds of printed pages, but did not mention spurious correlation. Yet although the debates treated here were, relatively speaking, asides, they are more revealing about interpretations of correlation than some of the noisier ones.

## 2. PEARSON ON CORRELATION AND CAUSATION

Pearson (1930, page 1) wrote of Galton and correlation: "Up to 1889 men of science had thought only in terms of causation, in future they were to admit another working category, that of correlation, and thus open to quantitative analysis wide fields of medical, psychological and sociological research." Edgeworth and Weldon took up correlation—theoretical and practical—before Pearson, but the change was achieved largely through his efforts.

Correlation was important for two of Pearson's interests: heredity and philosophy of science. Correlation was fundamental to the large program of theoretical and empirical work that Pearson described as "mathematical contributions to the theory of evolution." However, before he started researching into inheritance and developing the necessary statistical theory he had written about causation and scientific laws in *The Grammar of Science* (Pearson, 1892). Some of this work is relevant to correlation. There is a good brief account of Pearson's philosophy of science and its context in Passmore (1966, Chapter 14), while Norton (1975) treats the interplay between Pearson's philosophy of science and his work on heredity.

The first edition of *The Grammar of Science* (1892) predated Pearson's work on correlation, but the chapter on "Cause and Effect" is the same in all editions. Pearson (1892, page 155) treated causation as invariable temporal precedence: "Whenever a sequence of perceptions $D, E, F, G$ is invariably preceded by the perception $C \ldots$, $C$ is said to be a *cause* of $D, E, F, G$."

On this view it is easy to see correlation as a weaker form of causation. In his first paper applying correlation to meteorology, Pearson speaks of the "passage of correlation into causal relationship" when the correlation coefficient approaches 1; causation is perfect correlation:

> [I]f the unit $A$ be always preceded, accompanied or followed by $B$, and without $A$, $B$ does not take place, then we are accustomed to speak of a causal relationship between $A$ and $B$. (Pearson and Lee, 1897, page 459)

By the third edition of the *Grammar* (1910), Pearson was arguing that correlation was fundamental to the whole of science. There is a new section, "The Category of Association, as replacing Causation," where he urged:

> It is this conception of correlation between two occurrences embracing all relationshisp from absolute independence to complete dependence, which is the wider category by which we have to replace the old idea of causation. (Pearson, 1910, page 157)

Pearson even wondered whether "absolute dependence of a phenomenon on a single measurable cause" ever exists when "the refinement of observation is intense enough."

Pearson saw correlation as an innovation of epoch-making importance and he was unsparing in his criticism of shoddy work. According to Pearson and Elderton (1923, page 282):

> Such high correlations as arise from common growth or decline with time, when interpreted as causal or semicausal relationships, are in our opinion perfectly idle, indeed are only too apt to be mischievous, and we shall reach nothing, or less than nothing—knighthoods—by the investigation of them.

This is a nice example of Pearson's polemical style. Incidentally, in 1935 he refused a knighthood.

These ideas about correlation and causation provide the background to Pearson's thinking about spurious correlation. His complete position was that causation *is* correlation, or more precisely the limiting case of correlation, *except* when the correlation is spurious, when correlation is not causation. The *Grammar* provided the first half of the formula. The second half emerged as Pearson discovered cases of correlation that did not correspond to the kind of causation he had in mind in a particular scientific investigation. We shall see what these kinds were in Sections 3 and 9.

Pearson attached more and more importance to correlation, insisting ever more strongly on its centrality to all science, but he also found more and more situations in which correlation analysis was misleading. His discussions of spurious correlation were all occasional pieces but tying them together was the notion of a spurious correlation as one that did not indicate an "organic relationship." To judge from his scientific writings, the ultimate objective was to find evidence of organic relationships. For Pearson and Elderton (1923, pages 281–282) an organic relationship was a "causal or semicausal" relationship. Pearson used the term "organic relationship" as others used "law of nature" or "natural law"; for the case of Mill, see Aldrich (1992). The *Grammar* devoted a chapter to elucidating the notion of "law," but nowhere discussed what sets apart organic relationships from other regularities.

In much modern work on causation *hypothetical* outcomes or *hypothetical* interventions play an important role. The *Grammar* is hostile to such "unreal" constructions. Pearson does not belong in the ancestry of Rubin's (1974) construction (see Holland, 1986) or that of Simon (1954) (see Aldrich, 1989 and 1993) unless as an example of what they are trying to avoid.

## 3. COMMON CAUSES AND SPURIOUS CORRELATION

Of the possible types of organic relationship, Pearson most often invoked that of correlation through dependence on common causes. For Galton (1888, page 135) the correlation between two variables measures the extent to which they are governed by "common causes." Given his interests in heredity and physical anthropology, this was natural. Children resemble parents because they share some basic hereditary material. In the case of the correlation between the size of bodily organs there must be some organizing principle in the body behind such similarities. The length of one arm had no causal priority over the length of the other.

Pearson formalized the common cause scheme in his first paper (1896, pages 261–263) on correlation: the correlated observables are "determind by a great variety of *independent* contributory causes." This construction is an extension to the multivariate normal distribution of the derivation of the univariate normal from "elementary errors." Psychologists, following Spearman (1904), did much with common cause models. However, psychological factor analysis starts from the belief that there is one factor or at least only a few factors. For Pearson there were many factors and he had no interest in identifying them.

Pearson (1902, page 235) recalls he first became aware of a problem of spurious correlation in a study of "personal judgement" in taking astronomical readings. However, his paper on spurious correlation (Pearson, 1897) stays closer to the paradigm case of correlation by emphasizing the case of correlation body measurements.

Pearson's colleague, the biologist W. P. Weldon, had correlated the size of organs of crustaceans. His practice (e.g., Weldon, 1892 pages 3–4) was to express the measurements in terms of body length. Pearson saw a danger in this procedure, which he dramatized in the following way:

> For example, a quantity of bones are taken from an ossuarium, and are put together in groups, which are asserted to be those of individual skeletons. To test this a biologist takes the triplet femur, tibia, humerus and seeks the correlation between the indices femur/humerus and tibia/humerus. He might reasonably conclude that the correlation marked organic relationship, and believe that the bones had really been put together substantially in their individual grouping. As a matter of fact, since the coefficients of variation for femur, tibia and humerus are approxiamtely equal, there would be a correlation of about 0.4 to 0.5 between these indices had the bones been distributed absolutely at random. (Pearson, 1897)

The value of "about 0.4 to 0.5" came from a formula that Pearson developed for the correlation of $x_1/x_3$ and $x_2/x_3$ when $x_1$, $x_2$ and $x_3$ are independent random variables with equal coefficients of variation.

The bones of the same body are correlated in size through (unknown) causal agencies; the sum of these causal agencies constitutes the organic relationship the biologist is investigating. In Pearson's example of random groups of bones the correlation cannot be an expression of an organic relationship because none exists.

The example does not endanger the formal notion of correlation as a register of common causes because this could be saved by observing that the ratios *are* organically related by virtue of their common denominator. However, such "organic relationship" is not as interesting as that between bones of the same body and Pearson does not consider this way of saving the concept.

For Pearson, spurious correlation was a consequence of certain standard ways of processing data: in the case of this astronomical measurements, the values recorded by two observers might be referred to that obtained by a "standard observer" or in social or medical statistics, values might be deflated by population size. Pearson reacted to the problem in two ways: to be very wary of correlating ratios and, if forced to do so, to adopt as the point of no connection not 0, but some such value as 0.4.

In his earliest writings, Pearson (1902, page 235) found the source of spurious correlation in manipulations of the data: "[the] resemblance is due solely to the particular manipulation of the observations." We define spurious correlation to be a correlation which is produced by a process of arithmetic and not by any organic relationship among the quantities dealt with (1913, p. 150). In Pearson's later writings, when the field for spurious correlation had expanded, the fault of the statistician was less in inserting correlation than in failing to remove it.

## 4. YULE ON SPURIOUS CORRELATION DUE TO THE USE OF INDICES

Weldon, in a note attached to Pearson's paper (1897, page 498), accepted its argument and acknowledged that "the correlation due to indices forms 47 per cent of the observed value" of the correlation between the size of shrimp organs. Weldon did not explain why he used ratios in the first place. Neither he nor Pearson discussed why such manipulations are required nor how this relates to the way the correlation analysis should be carried out.

Yule used ratios extensively in his own correlation analysis: pauperism rates, birth rates and so forth. Nearly 15 years later, when Pearson, Lee and Elderton (1910) proposed a method for circumventing spurious correlation, Yule published a note which at once generalized Pearson's problem and took away its special sting for correlators of ratios. All correlation analysis faced problems of interpretation—they were not specially severe for correlation of ratios.

Yule had no quarrel with Pearson's distribution theory but he did dispute the inference that correlations with ratios were necessarily suspect. He (Yule, 1910, page 645) distinguished three cases that arise in correlating ratios: when "the causes the nature of which we wish to elucidate" influence directly the absolute magnitudes, when they influence the ratios and when "we have no direct knowledge of the mode of operation of the causes."

Yule agreed with Pearson that in the first case it would be misleading to correlate the ratios, but suggested that the appropriate procedure would be to correlate the absolute magnitudes. In the second case, it would be appropriate to correlate the ratios and the correlation between the absolute magnitudes would be misleading: Yule illustrated this case with the correlation between the death rate and sanitary conditions, where it would be inappropriate to use total numbers of deaths. He derived a formula analogous to Pearson's for the correlation between the absolute magnitudes when the ratios were uncorrelated.

For the third case, where it is not known how the causes operate, Yule argued as follows:

> This, as it seems to me, includes the case principally considered by Professor Pearson.... I concur with him in thinking that the interpretation of correlation between indices "is not free from obscurity," but, in my view, this obscurity is no less for correlation between absolute measurements, seeing that we have at present no real knowledge of the process of ontogeny to guide us. If the factors in the fertilised ovum, whatever they may be, which determine, for a given environment the ultimate form of the individual, determine directly the three distinct measurements $x_1$, $x_2$ and $x_3$, then certainly the correlation between the indices $x_1/x_3$ and $x_2/x_3$ is misleading, and cannot be regarded as an index to any relation between the germinal factors in question. If on the other hand, these germinal factors determine the two indices $z_1$ $(= x_1/x_3)$ and $z_2$ $(= x_2/x_3)$ and the measurement $x_3$ (two determinations of shape and a determination of the general size of the animal), then the correlations between the absolute measurements $x_1$ and $x_2$ are misleading. if the form of the animal is determined in some quite different way, both correlations become equally misleading. We have at present, so far as I am aware, no special reason for supposing that correlations between indices (organic indices) are more likely to be misleading than correlations between any other measures of shape or size. (Yule, 1910, page 647)

Yule's argument could obviously be extended to cover the case where there is an arbitrary system of variables: whatever the relationships or lack of relationships between the variables, correlated (or uncorrelated) variables could be manufactured out of them. What significance could be attached to the discovery of such correlations? They did not correspond to the laws of nature that biologists and others wished to uncover.

Pearson did not reply to Yule, but when Yule's article was noticed in a list of recent publications in

*Biometrika* (1913, page 158) it was with the comment, "The author makes a very curious blunder in criticising the theory of spurious correlation." Yule was hardly neglected because the same issue has Pearson and Heron's (1913) 150-odd page criticism of Yule's approach to association. In the next section we return to the years when Yule worked in Pearson's department, as his demonstrator and junior colleague.

## 5. YULE AND "CATENAS OF CAUSATION"

The interpretation of correlated variables as depending on common causes became common in psychology. A different interpretation became standard elsewhere: one variable causes (at least in part) the other. Yule was interested in socio-economic studies in which the correlated variables were either cause and effect or indicators of a cause and effect. Stigler (1986, pages 345–358) discussed Yule's work on correlation, concentrating on an aspect we do not consider here: how Yule took a scheme based on normality and applied it to nonnormal social data.

The new interpretation of correlation emerged in Yule's (1895, 1896 and 1899) studies of pauperism. These investigations into the correlates of pauperism were part of a debate about the effectiveness of different forms of poor relief. The main issue was whether the form of relief given (indoor, i.e., in the workhouse, or outdoor) affected the numbers of people receiving relief.

Yule was originally very skeptical about the causal import of correlation coefficients. He took a symmetrical view of the variables—though not one based on the common cause scheme. Thus he added a rider to the report that the rate of pauperism in a union (the unit of Poor Law administration) is positively correlated with the proportion of out-relief:

> This statement does not say either that the low mean proportion of out-relief is the cause of the lesser mean pauperism or vice versa: such terms seem best avoided where one is not dealing with a catena of causation at all. To use a simile, due I believe to Professor Marshall [1890, page 534] the case is like that of a lot of balls—say half a dozen—resting in a bowl. Then you cannot say that the position of ball No. 3 is the cause of the position of No. 5 nor the reverse. But the position of 3 is a function of the positions of all the others including 5; and the position of 5 is a function of the positions of all the others including 3: hence variations in the positions of the two balls will be correlated, and it is to this term that I prefer to adhere. To be quite clear, I do not mean simply that out-relief determines pauperism in one union and pauperism out-relief in another, so that you cannot say which is which in the average: but I mean that out-relief and pauperism mutually react in one and the same union. (Yule, 1895, page 605)

Yule did not take this notion of "mutual reaction" any further. An appropriate statistical model for Marshall's equilibrium economics was not fully developed until much later—the most important contribution is Haavelmo (1944). Aldrich (1993) discussed Haavelmo's analysis and gives references to the econometrics literature. Haavelmo's model combines elements of the Galton–Pearson scheme of correlation as expression of dependence on unobserved common causes and the scheme Yule did develop, namely, correlation as an expression of the degree to which one observed variable causes another.

In his next piece, Yule (1896) takes seriously the possibility that he had rejected, namely, that the form of administration could influence the rate of pauperism. Besides the possibility of a direct relationship, he also considered the possibility that the positive correlation could be explained by both variables being positively correlated with poverty itself. This possibility was closer to the common cause scheme, though the common cause was named and not confined to the limbo of unobserved elementary errors.

This study marked the debut of partial correlation which Yule had been working on. The full theory is set out in Yule (1897a, b) with an elegant reformulation in Yule (1907). Yule considered the partial correlation (or "net correlation") between pauperism and poor relief, holding poverty constant. When he found that the correlation persisted when allowance was made for the different levels of poverty in the different unions, his conclusion about the connection between the variables was that:

> This and some other work that I have done seems to indicate that no allowance for other factors will obscure the correlation between the two variables and to justify the assumption of a direct causal connection such as is usually postulated. (Yule, 1896, page 620)

The "catena of causation" was now justified; there was evidence of a "margin of immediately remediable pauperism."

The final paper in the series was "An Investigation into the Causes of Changes in Pauperism in

England" (Yule, 1899). This was a multiple regression study of a cross section of unions in which the dependent variable was the change over 10 years in the rate of pauperism and the independent variables were the change in the proportion receiving out-relief, the change in the proportion of the old and the change in population (an index to changes in prosperity).

Yule differentiated between the direct and indirect influences of the independent variables on the dependent variable. He took the multiple correlation exercise as clearing up the ambiguities of the total correlation coefficients, because the latter are compatible with several distinct hypotheses.

Partial correlations are not Yule's only resource in eliminating alternatives to the "catena of causation." He (Yule, 1899, page 275) noted that the correlation between change in pauperism and change in out-relief ratio could be interpreted in the opposite direction, however "it is difficult to imagine any causal relation between the two, such that pauperism should influence out-relief ratio."

## 6. SPURIOUS CORRELATION DUE TO THE "MIXING OF RACES"

In 1899, a second case of spurious correlation was discovered. Pearson was usually confident of discovering genuine correlations provided proper care was taken. Writing with Lee and Bramley-Moore (1899) he was less sure. They raised the possibility that the data being analyzed came from two distinct populations in each of which there is no correlation:

> We are thus forced to the conclusion that a mixture of heterogeneous groups, each of which exhibits in itself no organic correlation, will exhibit a greater or less amount of correlation. This correlation may properly be called spurious, yet as it is almost impossible to guarantee the absolute homogeneity of any community, our results for correlation are always liable to an error, the amount of which cannot be foretold. To those who persist on looking upon all correlation as cause and effect, the fact that correlation can be produced between two quite uncorrelated characters $A$ and $B$ by taking an artificial mixture of two closely allied races, must come as rather a shock. (Pearson, Lee and Bramley-Moore, 1899, page 278)

This discovery did not deflect Pearson for very long from placing great weight on correlations: he was himself the leader of "those who persist."

Yule adapted this example of heterogeneous groups to the discrete case. In 1900 Yule turned his attention to association between attributes and in particular to the problem of developing "some sort of 'coefficient of association' which should take the place of the 'coefficient of correlation' for continuous variables" (Yule, 1990). The measures that Yule devised—or the countermeasures of Pearson—are not particularly relevant to the present account of interpretations of correlation and association. The arguments about interpretation turn on independence or its absence, rather than on the measurement of association. The measures are discussed in Kendall and Stuart (1967, Chapter 33). We also will not follow the famous dispute with Pearson about Pearson's (1900) use of latent normal variables in the analysis of attributes. MacKenzie (1981, Chapter 7) has a nice account of this under the heading "Politics of the Contingency Table."

Yule's (1903) version of Pearson, Lee and Bramley-Moore's point appeared in the section of his paper on association on the "fallacies that may be caused by the mixing of distinct records." The source of the fallacy was the possibility that two attributes may be independent in subuniverses, but not in the universe.

Yule's (1903, page 134; 1922, page 50) favorite example involved drug testing. In line with his dominant interpretation of correlation and association, association of treatment and survival was taken as signifying the drug's efficacy. Suppose there is a different fatality rate for each sex when no drug is administered. If the female cases terminated fatally with greater frequency, but the drug was administered most often to males, then there would be created a "fictitious" association between the administration of the drug and recovery if the drug were completely ineffective.

In the "Simpson's paradox" (after Simpson, 1951) version, the attributes are associated in the subuniverses, but associated in the opposite direction in the universe. Yule's emphasis, like Pearson's, is on the mistaking of independence for dependence rather than on reversals of sign in dependence. Yule at least was aware of the Simpson possibility as shown by one of his exercises (Yule, 1922, pages 58 and 395) involving eye color in mothers, fathers and sons.

## 7. YULE ON "ILLUSORY ASSOCIATIONS"

Yule summarized his work on correlation and association in *An Introduction to the Theory of Statistics* (1911). Questions of interpretation are treated mainly in the more elementary context of association. Much of the chapter on partial association is concerned with "illusory association due to the association of two attributes with a third" (especially Yule, 1922, pages 48–51).

Yule's position—with his own terms in italics—is as follows: when an association leads to the inference of a *direct causal relation* where none exists it is *misleading*. The inferred causal relation is an *illusion*. It is a *fallacy* to interpret an association as if it were necessarily due to such a relation. There is room for confusion here. Yule applies "illusory"—his favorite term—to associations although the illusion is of a causal relation. There is a parallel with Pearson: in the phrase "spurious correlation," what is spurious is the inferred causal relation, or so Yule thought. However, for Pearson himself the causal relation is not something inferred from a correlation. It *is* a correlation or, at least, so it is described in the *Grammar*; see Section 2 herein.

Yule (1922, page 42) recognized that an association may be due to "fluctuations of sampling"—a possibility he worked hard in the time series context (Section 11 herein), but which he did not treat as problematic here. He applies the terms "illusory" and "misleading" to associations that provide ready support to hypotheses that turned out on further analysis to be false. The hypothesis is that the association is "due to the most obvious form of direct causation" (1922, page 48); the further analysis typically involves considering partial associations.

Yule's concern with alternative hypotheses continues from his work on pauperism. He gives several examples of fallacious reasoning, but there is no attempt to classify them. They include the inference to "$A$ causes $B$" when $A$ and $B$ are joint effects of $C$ but are otherwise independent and the inference to "$A$ has some effect on $C$" when $A$ has no effect on $C$ apart from its already well-established influence via $B$.

Yule emphasized the multiplicity of possible interpretations and the role of background knowledge when choosing one:

> The value of the coefficient may be consistent with some given hypothesis, but it may be equally consistent with others; and not only are care and judgement essential for the discussion of such possible hypotheses, but also a thorough knowledge of the facts in all other possible aspects. (Yule, 1922, page 191)

If there is a lesson in Yule's analyses of illusory associations, it is to look for homogeneous subuniverses, a concept that Lindley and Novick (1981) investigated using exchangeability. Partial associations based on plausible confounding factors seem likely to be useful. There are no techniques for adjusting for "misleading" correlations.

Yule's "misleading" becomes Simon's (1954, page 467) "spurious": $A$ and $B$ are not directly causally related, but are both effects of $C$. Unlike Simon, Yule did not consider how to discriminate between this case and another homogeneous subuniverse case when $A$ does not directly cause $B$ but acts *via* $C$; nor, of course, did Yule use any of the structural equation notions developed by econometricians and employed by Simon. Simon's change of terminology can be confusing: he writes that the "concept of spurious correlation" was "examined" by Yule (1922, page 215) reserved *that* name for Pearson's ratio case discussed previously in Section 3.

## 8. TIME SERIES CORRELATION ANALYSIS

When Pearson and Yule came to apply correlation to time series data they expected to find the same evidence of direct causal relationships and of dependence on common causes as in cross sections. What gave time series correlation its peculiar character was its attempt to accommodate the established practice of expressing a series as the sum of deterministic functions of time and error. Pearson and Yule disagreed over what the "time factor" signified and how it should be treated.

Yule was mainly concerned with movements and comovements of economic and demographic variables; see, for example, his investigation (Yule, 1906) of the causes of changes in marriage and birth rates. The comovements were often manifestations of the "trade cycle": "short period waves which form so conspicuous a feature of the curves of foreign trade, the marriage rate, prices, employment and so forth" (Yule, 1915, page 305).

Yule (1909, pages 725–728) advocated a form of correlating adjusted time series developed by his friend and collaborator R. H. Hooker. Hooker examined whether the marriage rate "fluctuates according to the general prosperity of the country, as indicated by the course of trade." He found a very low correlation for the period 1861–1895:

> In this case the difference in the general movement of the two curves has completely overshadowed the minor oscillations; whereas it is only to these latter that we refer when, in ordinary parlance, we speak of the marriage-rate as being dependent upon trade. (Hooker, 1901, page 486)

He hoped to get at these "minor oscillations" by basing his correlations, not on deviations from the mean of the whole period, but on deviations from a trend. When the series were centered on moving average trends, the correlation was raised substantially.

Hooker introduced another data-transformation device for emphasizing a particular period. He (Hooker, 1905, page 703) proposed the correlation of the first differences of the series as a measure of the similarity of the "shorter rapid changes (with no apparent periodicity)". He did not deprecate the correlation between the unadjusted data: this correlation was the "most suitable test of 'secular' independence" (Hooker, 1905, page 703).

By the "dependence" of the marriage rate on general prosperity, Hooker meant direct causal dependence as in Yule's work on pauperism. Yule did not focus so exclusively on direct dependence. He (Yule, 1921a, page 503) wrote "the correlation between [oscillatory] movements may arise either because the one is causally dependent on the other, or because both are functions of the time, i.e., of some third variable or group of variables, on which both are causally dependent." Yule did not apply the term "illusory" to the second situation.

Techniques like Hooker's—with what seemed to be differences of detail only—were used in Pearson's laboratory. Thus Cave-Brown-Cave (1904) worked with correlations of differenced temperatures. Yet there developed a spectacular divergence of view about the point of such calculations. This was brought out by some work of Student's that Pearson promoted. This is discussed in the next section. E. S. Pearson (1990, pages 29–34) gives a fuller account.

## 9. SPURIOUS CORRELATION "DUE TO POSITION IN TIME"

Student's "The Elimination of Spurious Correlation due to Position in Time or Space" (1914) claims to be an extension of Hooker's method of correlating differences. According to Student:

> [The method] helps to show, whether, for example there really is a close connexion between the female cancer death rate and the quantity of imported apples consumed per head! (Student, 1914, page 179)

Student's method, which Cave and Pearson (1914) called the variate difference correlation method, was a method of estimating a parameter in an explicit model. For our purposes the method is less important than the model. The variables $x$ and $y$, for example, the female cancer death rate and the quantity of imported apples consumed per head, are decomposed into polynomial time trends and terms that are "independent of time,"

$$x_t = \alpha_1 t + \cdots + \alpha_n t^n + X_t,$$
$$y_t = \beta_1 t + \cdots + \beta_n t^n + Y_t,$$

where $X_t$ and $Y_t$ are independent and identically distributed random variables with contemporaneous correlation $r_{XY}$, but are otherwise uncorrelated.

The components $X$ and $Y$ are "masked by the time effect." The correlation between the variables $r_{xy}$ may give a misleading account of the relation between them. It is the correlation $r_{XY}$ which indicates whether there is "really" a "close connexion" between $x$ and $y$.

The Yule of "illusory associations" should have liked this model because his partialling out corresponds to the removal of the time component. However, his paper "On the Time-Correlation Problem, with Special Reference to the Variate-Difference Correlation method" (1921a) was an attack of Pearsonian proportions on the model, the interpretation of $r_{XY}$ and the estimation method.

Yule emphasized how Student's "generalised method" deviated from Hooker's:

> [T]he problem is not to isolate random residuals but oscillations of different durations, and unless the generalised method can be given some meaning in terms of oscillations it is not easy to see what purpose it can serve. (Yule, 1921a, page 504)

Yule (1921a, page 504) identified "random residuals" with "errors of observation, errors due to the 'rounding off' of index numbers and the like, fluctuations of sampling and analogous variations." They were not interesting and Yule anticipated that $r_{XY}$ would be zero anyway.

In the discussion of Yule's paper, Greenwood defended Student's method. He (page 528) gave an example of a problem involving death rates from different diseases for which the method would be useful:

> [W]e do desire to eliminate both secular and oscillatory effects, and, if those effects can be eliminated, what remains is surely of the greatest importance. Indeed, so far as the statistical study of aetiology is concerned... I should be disposed to... say that the "problem is to isolate random residuals."

Greenwood puts the trend and periodic components aside because their action is understood, not because they generate spurious correlation. His example differs from the cancer rate example in Student's paper in that Greenwood assumes much greater background knowledge. Student's exercises are shots in the dark with the variate difference method sorting the hits from the misses. For Greenwood, the applications of the method were not like this: there was some independent ground for supposing that a relationship existed.

## 10. NEW MODELS FOR TIME SERIES

Disagreement might persist over which was signal, which was noise, which was baby, which was bathwater, . . . , but there was a model encompassing the views of all parties to the debate on the variate difference method:

$$x_t = \text{cycles} + \text{trend} + X_t,$$
$$y_t = \text{cycles} + \text{trend} + Y_t.$$

However, Yule soon moved away from this scheme and developed a new model, or collection of models. These were new univariate models. He developed no new bivariate model although this did not stop him from discussing correlation.

Yule not only moved from decomposing series into deterministic components, but he found a new respect for "random residuals." He (Yule, 1927) proposed the second order autoregressive process for the sun-spot series—a provocative choice for this had often been studied with an eye for hidden periodicities.

Yule (1921a) had used random series as test series for considering the effects of differencing. The idea behind Student's *method* was that the value of the correlation coefficient between the $n$th differences of $x$ and $y$ would equal $r_{XY}$. Yule investigated the effects of differencing on series constructed from trigonometric functions and random terms. His use of random series as a test series and analysis of the serial correlation in the resulting moving average pointed to a new type of time series correlation analysis. He started to build processes out of random components.

## 11. "NONSENSE-CORRELATIONS" BETWEEN TIME SERIES

Yule (1926) asked "Why do we sometimes get Nonsense-Correlations between Time-Series?," for example, a correlation of +0.9512 between the mortality rate for the years 1866–1911 and the ratio of Church of England marriages to all marriages. He rejected the explanation he had favored in other circumstances (see Section 9):

> Now it has been said that to interpret such correlations as implying causation is to ignore the common influence of the time factor . . . . I cannot regard time per se as a causal factor; and the words only suggest that there is some third quantity varying with the time to which the changes in both the observed variables are due . . . . But what one feels about such a correlation is, not that it must be interpreted in terms of some very indirect catena of causation, but that it has no meaning at all; that in non-technical terms it is simply a fluke, and if we had or could have experience of the two variables over a much longer period of time we could not find any appreciable correlation between tem. But to argue like this is, in technical terms to imply that the observed correlation is only a fluctuation of sampling, whatever the ordinary formula, for the standard error may seem to imply. (Yule, 1926, page 4)

The "ordinary formula" for the case of random sampling from the normal distribution had been developed by Pearson (1896) and Pearson and Filon (1898). Student (1908) gave the exact distribution of the correlation in the null case. Yule realized that this sampling theory could be misleading for serially correlated observations. He considered cases where the correlation coefficient was more dispersed than in the random case. Edgeworth in the discussion of Hooker (1907, page 44) had insisted that the $\sqrt{n}$ in the denominator of the formula for the probable error of the correlation should be taken "cum grano." It was to this time series paper and the discussion around it that Student (1908, page 302) had referred for expert opinion on the "limit of significance" of the correlation coefficient in small samples.

Yule was after more spectacular cases of the "ordinary formula" being wrong, to explain the correlation between the mortality rate and the proportion of Church of England marriages. He asked the following question:

> What characterisics must two empirical series possess in order that small random samples, taken from them . . . may tend to give a U-shaped frequency distribution for the resultant correlations? (Yule, 1926, page 14)

On the basis of some very heuristic reasoning and of evidence from sampling experiments, he suggested a type of series with the required properties: a "conjunct series with conjunct differences," that is, a series the second differences of which are random.

In his discussion of the paper, Edgeworth (page 66) missed the sampling point and went back to the common cause interpretation which Yule was explicitly rejecting. Edgeworth referred to Mill's "method of concomitant variations." Mill (1843, page 401) wrote, "it by no means follows, when two phenomena accompany each other in their variations, that the one is cause and the other effect. The same thing may, and indeed must happen, supposing them to be different effects of a common cause."

Yule did not propose a new bivariate model. His analysis of the distribution of the correlation under the null hypothesis of independence did not require one. It is not clear whether he was trying to present a new approach to time series correlations or just clearing up some paradoxical cases. His suggestion of considering the internal structure of each series to see whether a nonsense correlation was possible was not taken up. The work remained isolated until Hendry (1986) related it to Granger and Newbold's (1974) work on "spurious regressions" involving independent nonstationary series and to more recent work on cointegration.

## 12. DISCUSSION

Neither Yule nor Pearson gave a systematic account of correlations and their pathology. Yet there was a pattern in the way they treated the pathological cases. Here we try to identify it.

In interpreting Yule it is helpful to employ a distinction made by some modern writers, for example, Koopmans and Reiersøl (1950) and Cox (1958), between "scientific inference" and "statistical inference." Statistical inference deals with inference from sample to population while scientific inference deals with the interpretation of the population in terms of a theoretical structure. Koopmans and Reiersøl drew principally on their experience in econometrics and factor analysis (see Morgan, 1990, Chapter 6, and Aldrich, 1994), but their categories can be applied to correlation: inferring correlations versus understanding what they signify.

R. A. Fisher's "The correlation between relatives on the supposition of Mendelian inheritance" (1918) was the most ambitious piece of "scientific inference" using correlations in the era studied here. Fisher showed how the biometricians' correlations could be understood in terms of Mendelian genetics. Both Pearson and Yule [see, e.g., Pearson (1909) and Yule (1907a)] contributed to this project. Pearson contributed the estimates of population correlations to be rationalized but he also tried to show that biometric findings could not be explained in Mendelian terms. Nevertheless, as Provine (1971) and Norton (1975) show, his work—as well as Yule's more constructive efforts—was important in suggesting how the reconciliation might be achieved. All this work involved delicate interpretations of correlations.

Another field involving sophisticated interpretations of correlations—also based on "common causes"—was psychological factor analysis. Here there was a long debate between C. Spearman and G. H. Thomson on whether the single factor theory provided the only explanation of the observed correlations. Yule (1921a, p. 105) commented on the debate when he reviewed Thomson's book: "From the statistical standpoint Dr. Spearman's explanation seems to me to be by far the simplest, but the judgment as to its validity must be based on other grounds."

Besides these highly organized schemes for interpreting correlations, there was an exploratory correlation analysis which treated questions like, *is* there a "connexion between the female cancer death rate and the quantity of imported apples consumed per head?" In this field questions of the genuineness of correlations thrived.

Yule distinguished two ways in which correlation inferences could go wrong. In "illusory" correlations the errors are of interpretation—misconstruing the nature of a catena of causation—and as such faulty scientific inferences. The correlation estimates are taken as satisfactory. He wrote in his textbook:

> Any interpretation of the meaning of an association is necessarily hypothetical, and the number of possible alternative hypotheses is in general considerable. (Yule, 1911, p. 42)

Usually, though, Yule had in mind a best interpretation based on a priori information—most often of a direct causal connection—but his criticism of Pearson's "spurious correlation due to the use of ratios" recognized that there might not be enough structural knowledge to interpret the correlation. Fisher (1925, p. 133) put this view strongly—although in 1916 he had given an enthusiastic welcome to the variate-difference method—:

> If we choose . . . a group of social phenomena with no antecedent knowledge of the causation or absence of causation among

> them, then the calculation of correlation coefficients, total or partial, will not advance us a step towards evaluating the importance of the causes at work.

In econometrics the view became a dogma.

Correlation inferences could also go wrong in a different way, illustrated by the "nonsense-correlations" that can arise with independent time series. Incorrect conclusions were the result of faulty statistical inference based on the use of the wrong sampling theory. Here the statistician has a special responsibility, unlike the case of faulty scientific inference about which the statistician has nothing special to say. Misuse of sampling theory allows flukes to pass for regularities.

Pearson did not separate the estimation of correlations from their interpretation so sharply as Yule. For the author of the *Grammar*, there was no causation "behind" correlation since causation *is* correlation. For the scientist or applied statistician there was always a particular kind of organic relationship, a specific causal structure, relevant in any given context. Genuine correlation is associated with this relationship. Yule was always more mindful of alternative possibilities even if he rejected them.

Pearson also held that observed correlations may be compounded of spurious and genuine correlation; spurious correlation is something to be "eliminated." In modern terms an observed correlation is a more or less biased, or inconsistent, estimate of a causally relevant correlation coefficient—the extent of the bias measuring the spuriousness of the correlation.

A century after the problem of disentangling genuine from spurious correlation was first posed there is much more technique available for treating it. Theories of causation, such as those of Simon or Spirtes, Glymour and Schienes, identify causally relevant quantities; theories of estimation treat the best way of estimating any desired quantity; and theories of misspecification treat the consequences for inference of using an inappropriate model. These theories go deeper than Pearson and Yule yet their examples still hold their place.

## ACKNOWLEDGMENTS

I am grateful to Professor Robert Kass and two referees for their suggestions.

## REFERENCES


ALDRICH, J. (1989). Autonomy. *Oxford Economic Papers* **41** 15–34.

ALDRICH, J. (1992). Probability, statistics and political economy in Mill's *Logic*. Unpublished manuscript.

ALDRICH, J. (1993). Cowles exogeneity and CORE exogeneity. Discussion Paper 9308, Dept. Economics, Southampton Univ.

ALDRICH, J. (1994). Haavelmo's identification theory. *Econometric Theory* **10** 198–219.

ALDRICH, J. (1995). R. A. Fisher and the Making of Maximum Likelihood 1912–22. Discussion Paper 9504, Dept. Economics, Southampton Univ.

CARTWRIGHT, N. (1983). *How the Laws of Physics Lie*. Oxford Univ. Press.

CAVE, B. M. and PEARSON, K. (1914). Numerical illustrations of the variate-difference correlation method. *Biometrika* **10** 340–355.

CAVE-BROWNE-CAVE, F. E. (1904). On the influence of the time factor on the correlation between the barometric heights at stations more than 1000 miles apart. *Proc. Roy. Soc. London. Ser. A* **74** 403–411.

COX, D. R. (1958). Some problems connected with statistical inference. *Ann. Math. Statist.* **29** 357–372.

EISENHART, C. (1974). Karl Pearson. *Dictionary of Scientific Biography* **10** 447–473.

FISHER, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika* **10** 507–521.

FISHER, R. A. (1916). "Biometrika." *Eugenics Review* **8** 62–64.

FISHER, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* **52** 399–433.

FISHER, R. A. (1925). *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.

GALTON, F. (1888). Co-relations and their measurement. *Proc. Roy. Soc. London Ser. A* **45** 135–145.

GALTON, F. (1890). Kinship and correlation. *North American Review*. [Reprinted with an introduction by S. M. Stigler in *Statist. Sci.* (1989) **4** 73–86.]

GRANGER, C. W. J. and NEWBOLD, P. (1974). Spurious regressions in econometrics. *J. Econometrics* **2** 111–120.

GREENWOOD, M. and YULE, G. U. (1915). The statistics of anti-typhoid and anti-cholera inoculations, and the interpretation of such statistics in general. *Proceedings of the Royal Society of Medicine (Epidemiology)* **8** 113–190.

HAAVELMO, T. (1944). The probability approach in econometrics. *Econometrica (Suppl.)* **12** 1–112.

HENDRY, D. F. (1986). Economic modelling with cointegrated variables: an overview. *Oxford Bulletin of Economics and Statistics* **48** 201–212.

HOLLAND, P. M. (1986). Statistics and causal inference (with discussion). *J. Amer. Statist. Assoc.* **81** 945–971.

HOOKER, R. H. (1901). On the correlation of the marriage-rate with trade. *J. Roy. Statist. Soc. Ser. A* **64** 485–492.

HOOKER, R. H. (1905). On the correlation of successive observations illustrated by corn prices. *J. Roy. Statist. Soc. Ser. A* **68** 696–703.

HOOKER, R. H. (1907). The correlation of the weather and crops (with discussion). *J. Roy. Statist. Soc. Ser. A* **70** 1–51.

KENDALL, M. G. and STUART, A. (1967). *The Advanced Theory of Statistics*, 2nd ed. **2**. Griffin, London.

KOOPMANS, T. C. and REIERSØL, O. (1950). The identification of structural characteristics. *Annals of Mathematical Statistics* **21** 165–181.

KRONMAL, R. A. (1993). Spurious correlation and the fallacy of the ratio standard revisited. *J. Roy. Statist. Soc. Ser. A* **156** 379–392.

LINDLEY, D. V. and NOVICK, M. R. (1981). The role of exchangeability in inference. *Ann. Statist.* **9** 45–58.

MARSHALL, A. (1890). *The Principles of Economics*. Macmillan, London.

MILL, J. S. (1843). *A System of Logic: Ratiocinative and Inductive*. [Reprinted in 1973 (J. M. Robson, ed.), Univ. Toronto Press.]

MORGAN, M. S. (1990). *The History of Econometric Ideas*. Cambridge Univ. Press.

NORTON, B. J. (1975). Biology and philosophy: the methodological foundations of biometry. *Journal of the History of Biology* **8** 85–93.

PASSMORE, J. (1966). *A Hundred Years of Philosophy*. Penguin, Harmondsworth, Middlesex.

PEARSON, E. S. (1936). Karl Pearson: an appreciation of some aspects of his life and work. Part I. *Biometrika* **28** 193–257.

PEARSON, E. S. (1938). Karl Pearson: an appreciation of some aspects of his life and work. Part II. *Biometrika* **29** 161–247.

PEARSON, E. S. (1990). *"Student," A Statistical Biography of William Sealy Gosset* (R. L. Plackett and G. A. Barnard, eds.). Oxford Univ. Press.

PEARSON, K. (1892). *The Grammar of Science*. Walter Scott, London.

PEARSON, K. (1896). Mathematical contributions to the theory of evolution. III. Regression, heredity and panmixia. *Philos. Trans. Roy. Soc. London Ser. A* **187** 253–318.

PEARSON, K. (1897). On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proc. Roy. Soc. London Ser. A* **60** 489–498.

PEARSON, K. (1900). On the correlation of characters not quantitatively measurable. *Philos. Trans. Roy. Soc. Ser. A* **195** 1–47.

PEARSON, K. (1902). On the mathematical theory of errors of judgement, with special reference to the personal equation. *Philos. Trans. Roy. Soc. Ser. A* **198** 235–299.

PEARSON, K. (1909). On the ancestral gametic correlations of a Mendelian population mating at random. *Proc. Roy. Soc. London Ser. A* **81** 225–229.

PEARSON, K. (1910). *The Grammar of Science*, 3rd ed. Black, Edinburgh.

PEARSON, K. (1913). Lectures on Applied Statistics, notes taken by L. Isserlis. Item 84/2 in Pearson collection, Univ. College, London.

PEARSON, K. (1930). *The Life, Letters and Labours of Francis Galton*, Vol. IIIA. Cambridge Univ. Press.

PEARSON, K. and CAVE-BROWNE-CAVE, F. E. (1902). On the correlation between the barometric height at stations on the eastern side of the Atlantic. *Proc. Roy. Soc. London Ser. A* **70** 465–470.

PEARSON, K. and ELDERTON, E. M. (1923). On the variate difference method. *Biometrika* **14** 281–310.

PEARSON, K. and FILON, L. M. G. (1898). Mathematical contributions to the theory of evolution IV. On the probable errors of frequency constants and on the influence of random selection on variation and correlation. *Philos. Trans. Roy. Soc. Ser. A* **191** 229–311.

PEARSON, K. and HERON, D. (1913). On theories of association. *Biometrika* **9** 159–315.

PEARSON, K. and LEE, A. (1897). On the distribution of the frequency (variation and correlation) of the barometric heights at divers stations. *Philos. Trans. Roy. Soc. London Ser. A* **190** 423–470.

PEARSON, K., LEE, A. and BRAMLEY-MOORE, L. (1899). Genetic (reproductive) selection: inheritance of fertility in man, and of fecundity in thoroughbred racehorses. *Philos. Trans. Roy. Soc. London Ser. A* **192** 257–330.

PEARSON, K., LEE, A. and ELDERTON, E. M. (1910). On the correlation of death rates. *J. Roy. Statist. Soc. Ser. A* **73** 534–539.

PORTER, T. M. (1986). *The Rise of Statistical Thinking*. Princeton Univ. Press.

PROVINE, W. B. (1971). *The Origin of Theoretical Population Genetics*. Univ. Chicago Press.

RUBIN, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* **66** 688–701.

SIMON, H. A. (1954). Spurious correlation: a causal interpretation. *J. Amer. Statist. Assoc.* **49** 467–492.

SIMPSON, E. H. (1951). The interpretation of interaction in contingency tables. *J. Roy. Statist. Soc. Ser. B* **13** 238–241.

SPEARMAN, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology* **15** 201–293.

SPIRTES, P., GLYMOUR, C. and SCHIENES, R. (1993). *Causation, Prediction and Search*. Springer, New York.

STIGLER, S. M. (1986). *The History of Statistics*. Harvard Univ. Press.

STUDENT (1908). Probable error of a correlation coefficient. *Biometrika* **6** 302–310.

STUDENT (1914). The elimination of spurious correlation due to position in time or space. *Biometrika* **10** 179–180.

WELDON, F. R. (1892). Certain correlated variations in *Crangon vulgaris*. *Proc. Roy. Soc. London Ser. A* **51** 2–21.

YATES, F. (1952). George Udny Yule, 1871–1951. *Obituary Notices of Fellows of the Royal Society* **8** 309–323.

YULE, G. U. (1895). On the correlation of total pauperism with proportion of out-relief. I. All ages. *Economic Journal* **5** 603–608.

YULE, G. U. (1896). On the correlation of total pauperism with proportion of out-relief. II. Males over 65. *Economic Journal* **6** 613–623.

YULE, G. U. (1897a). On the significance of Bravais' formulae for regression, &c., in the case of skew correlation. *Proc. Roy. Soc. London Ser. A* **60** 477–489.

YULE, G. U. (1897b). On the theory of correlation. *J. Roy. Statist. Soc. Ser. A* **60** 812–854.

YULE, G. U. (1899). An investigation into the causes of changes in pauperism in England, etc. *J. Roy. Statist. Soc. Ser. A* **62** 249–296.

YULE, G. U. (1900). On the association of attributes in statistics: with illustrations from the material of the Childhood Society, &c. *Philos. Trans. Roy. Soc. London Ser. A* **194** 257–319.

YULE, G. U. (1903). Notes on the theory of association of attributes in statistics. *Biometrika* **2** 121–134.

YULE, G. U. (1906). On the changes in the marriage- and birthrates in England and Wales during the past half century; with an inquiry as to their probable causes (with discussion). *J. Roy. Statist. Soc. Ser. A* **69** 88–147.

YULE, G. U. (1907a). On the theory of correlation for any number of variables, treated by a new system of notation. *Philos. Trans. Roy. Soc. London Ser. A* **79** 182–193.

YULE, G. U. (1907b). On the theory of inheritance of quantitative compound characters on the basis of Mendel's laws—a preliminary note. In *Report of the Third International Conference on Genetics* 140–142. Spottiswoode, London.

YULE, G. U. (1909). The applications of the method of correlation to social and economic statistics. *J. Roy. Statist. Soc. Ser. A* **72** 721–730.

YULE, G. U. (1910). On the interpretation of correlations between indices or ratios. *J. Roy. Statist. Soc. Ser. A* **73** 644–647.

YULE, G. U. (1911). *An Introduction to the Theory of Statistics*, 1st ed. Griffin, London. [All page references are to the sixth edition (1922).]

Yule, G. U. (1915). Review of H. L. Moore's *Economic Cycles*. *J. Roy. Statist. Soc. Ser. A* **78** 302–305.

Yule, G. U. (1921a). On the time-correlation problem, with special reference to the variate-difference correlation method. *J. Roy. Statist. Soc. Ser. A* **84** 497–526.

Yule, G. U. (1921b). Critical notice of "The Essentials of Mental Measurement" by W. Brown & G. H. Thomson. *British Journal of Psychology* **12** 100–107.

Yule, G. U. (1922). *An Introduction to the Theory of Statistics*, 6th ed. Griffin, London. (First edition, 1911.)

Yule, G. U. (1926). Why do we sometimes get nonsense correlations between time-series? A study in sampling and the nature of time-series. *J. Roy. Statist. Soc. Ser. A* **89** 1–69.

Yule, G. U. (1927). On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philos. Trans. Roy. Soc. London Ser. A* **226** 267–298.

Yule, G. U. (1936). Karl Pearson 1857–1936. *Obituary Notices of Fellows of the Royal Society* **2** 73–104.