

The Scientific Work of Robert Charles Geary

JOHN E. SPENCER

“Go on, Mr. Pratt,” says Mrs. Sampson, “Them ideas is so original and soothing. I think statistics are just as lovely as they can be.”
(From *The Handbook of Hymen* by O. Henry)

ROY GEARY is eighty this year and for almost half a century he has enjoyed a major international reputation in mathematical statistics. In the course of a long and varied career he has been Director of both the Central Statistics Office and The Economic and Social Research Institute in Dublin, Chief of the UN's National Accounts Branch, and has held posts in Universities in Britain and the USA. He is an Honorary Fellow of several learned societies* and has been a member of council both of the International Association for Research in Income and Wealth and of the Econometric Society. He has to date published almost one hundred papers mainly in the field of mathematical statistics, but he has also contributed to economic theory, applied economics, economic statistics and demography.

His versatility was neatly summarised by Sir Maurice Kendall who gave the Sixth Geary Lecture in 1973: “My distinguished predecessor in contributing to

*The list is impressive:

Past President, Statistical and Social Inquiry Society of Ireland.

Past President, International Statistical Institute.

Vice-President, Royal Irish Academy.

Honorary Fellow, Royal Statistical Society.

Honorary Fellow, American Statistical Association.

Honorary Fellow, Institute of Mathematical Statistics.

Honorary Fellow, Institute of Statistics.

Past Member of Council, International Association for Research in Income and Wealth.

Past Member of Council, Econometric Society—(*Editor.*)

this series of lectures, Jan Tinbergen, included in the title of his lecture the word 'interdisciplinary'. I follow him in spirit, for the problem of forecasting is truly interdisciplinary, calling as it does on the combined skills of the economist, the statistician and the mathematician as well as the commonsense of the practitioner. And I cannot think of any more suitable name in whose honour a lecture of this kind should be given than that of Roy Geary, himself an interdisciplinarian if ever there was one, equally at home in all these subjects, and among his many distinctions possessing one which I think is unique, that of being the only former head of a Government Central Statistical Office whose name is attached to a mathematical theorem, has acted at the Abbey Theatre and has been offered a job as a professional footballer."

As a tribute to the greatest Irish statistician of our time, I have prepared an account of most of his major contributions to statistical theory and methods. Accordingly, this paper will only give a partial picture of his work since it omits his many contributions in applied statistics and economics and does not attempt to appraise his considerable influence on the methodology of national accounting. However, his contributions to theory probably comprise the most important part of his work and it is, perhaps, the part of which he is most proud.

"Mathematics is . . . the sublime art . . . Mathematica remains a bit aloof from, and disdainful of the business of life. She belongs to the Arts. She is proudly her own *raison d'être*" (Geary, 1964).

It is always of interest to read the first paper of a man who subsequently evolves into a great scholar. Geary's first paper (1925) is no exception, showing fine technical ability, a natural flair for stochastic problems and a painstaking care for detail, qualities which were to manifest themselves throughout his later work. This early paper also provides an excellent example of the power of good theory when applied to real problems; here the problem concerned Irish agricultural statistics which had been thrown into some confusion during political troubles of the time. The theoretical result is as follows. Let $m+u_i, i = 1 \dots N$ be the values of N elements in year one and $m'+u'_i, i = 1 \dots N$ be the values in year two, where m and m' are the means in the two years. It is desired to measure the ratio of true means m'/m by taking a random sample of n elements. The ratio as estimated by $\sum_1^n (m'+u'_i) / \sum_1^n (m+u_i)$ is shown to be approximately normally distributed for large n and N with a mean of m'/m . Geary also computes the variance and applies the result successfully to agricultural data.

This interest in finding the density of a ratio continued and five years later the classic Geary (1930b) paper appeared. Let x_1 and x_2 be two jointly distributed normal variables with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2 respectively, with correlation r . The difficult technical problem of finding the exact sampling distribution of $z = x_1/x_2$ was solved on the hypothesis that μ_2 was large relative to σ_2 so that the range of x_2 was effectively positive. On this assumption Geary proved that the ratio $(\mu_2 z - \mu_1) / \sqrt{(\sigma_2^2 z^2 - 2r\sigma_1\sigma_2 z + \sigma_1^2)}$ was distributed $N(0, 1)$. The result

in its original form is still quoted today. (See, for example, Scadding (1973) and the survey article Marsaglia (1965).) An expression for the density of z where x_1 and x_2 are independent but not necessarily normal variates was discovered by Cramer (1937) for the case where x_2 is non-negative with finite mean. Writing $\Phi_1(t)$ and $\Phi_2(t)$ for the characteristic functions of x_1 and x_2 , the density of z , $f(z)$, if it exists was found to be

$$f(z) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \Phi_1(t) \Phi_2'(-tu) dt,$$

provided the integral converges. Geary (1944b) generalised this result to the case of non-independent x_1 and x_2 with joint characteristic function $\Phi(t_1, t_2)$ and the same condition on x_2 . The generalised result is

$$f(z) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \delta\Phi(t_1, t_2)/\delta t_2 dt$$

where the partial derivative is evaluated at $t_2 = -t_1u$.

That Geary is an expert in the use of cumulants is apparent from his earliest papers. In order to explain some of this early work, recall that the cumulants of a random variable are defined from the formal expansion of the characteristic function (cf), $\Phi(t)$.

Writing $\Phi(t) = \int e^{itx} dF$ where F is the distribution function of the random variable x , suppose that $\Phi(t)$ can be expressed as $\exp(\sum_1^{\infty} \kappa_j \Theta^j/j!)$ where $\Theta = it$. The cumulants are then defined as the κ_j in this expansion and $\log \Phi(t)$ can clearly be regarded as the cumulant generating function since differentiating p times with respect to Θ and setting $\Theta = 0$ in the result will give the p 'th cumulant κ_p .

Similarly we can define the joint cumulants of a joint distribution from the equation $\Phi(t_1, t_2) = \exp(\sum_0^{\infty} \sum_0^{\infty} \kappa_{rs} \Theta_1^r \Theta_2^s/r! s!)$ with $\log \Phi(t_1, t_2)$ again acting as the (joint) generating function.

Cumulants can be equivalently defined in terms of the (assumed to exist) moments about the origin of a random variable (or of jointly distributed random variables) through the identity $\Phi(t) = \sum_0^{\infty} \mu_j' \Theta^j/j!$. The κ 's can then be expressed as functions of the μ 's (and vice versa) enabling sample cumulants to be defined through the replacement of population raw moments with sample raw moments.

In two brilliant papers (1942b, 1943c), Geary applied cumulant theory to the estimation of the coefficients of the linear relationship $\beta_1 X_1 + \dots + \beta_k X_k = 0$ where the X 's are measured with error so that the observables are $x_j = X_j + u_j$. These papers are ingenious and technically highly instructive. In order to simplify exposition, suppose that all variables have zero means that the u 's are independent of each other and of the X 's, and that $k = 2$. Thus $\beta_1 X_1 + \beta_2 X_2 = 0$ or $X_2 = \beta X_1$ where $\beta = -\beta_1/\beta_2$.

The joint characteristic function of X_1 and X_2 is $\Phi(t_1, t_2)$.

Clearly $\beta_1 \delta\Phi/\delta\Theta_1 + \beta_2 \delta\Phi/\delta\Theta_2 = E(\beta_1 X_1 + \beta_2 X_2) \exp(\Theta_1 X_1 + \Theta_2 X_2) = 0$.

Hence $\sum \beta_j \delta \log \Phi/\delta \Theta_j = 0$ for all Θ_1, Θ_2 .

Differentiating the latter identity (in Θ) p_1 times with respect to Θ_1 and p_2 times with respect to Θ_2 , we have

$$\beta_1 \kappa_{p_1+1, p_1} + \beta_2 \kappa_{p_1, p_1+1} = 0 \quad \text{so that}$$

$$\beta = (\kappa_{p_1, p_2+1}) / (\kappa_{p_1+1, p_2}) \quad \text{provided the denominator is non zero.}$$

Now the joint characteristic function of the observed x 's is $E \exp(\Theta_1 x_1 + \Theta_2 x_2)$ which can be written using the postulated relationships and independence assumptions as $E \exp(\Theta_1 X_1 + \Theta_2 X_2)$. $E \exp(\Theta_1 u_1 + \Theta_2 u_2)$. The next step is fundamental.

Write $\bar{\Phi}(t_1, t_2)$ as the cf of the x 's and $\bar{\bar{\Phi}}(t_1, t_2)$ as the cf of the u 's with analogous notation for the cumulants.

$$\begin{aligned} \text{Then } \log \bar{\Phi}(t_1, t_2) &= \sum \sum \bar{\kappa}_{rs} \Theta_1^r \Theta_2^s / (r! s!) \quad (\text{by definition of cumulants}) \\ &= \log \Phi(t_1, t_2) + \log \bar{\bar{\Phi}}(t_1, t_2) \quad (\text{by the preceding paragraph}). \end{aligned}$$

$$\begin{aligned} \text{But } \log \bar{\bar{\Phi}} &= \sum \sum \bar{\bar{\kappa}}_{rs} \Theta_1^r \Theta_2^s / r! s! \\ &= \log(\text{cf of } u_1) + \log(\text{cf of } u_2) \quad (\text{by independence of } u_1 \text{ and } u_2) \\ &= \sum_1^\infty \bar{\bar{\kappa}}_{r0} \Theta_1^r / r! + \sum_1^\infty \bar{\bar{\kappa}}_{0s} \Theta_2^s / s! \end{aligned}$$

so that $\bar{\bar{\kappa}}_{rs} = 0$ for $rs \neq 0$.

Hence, $\kappa_{rs} = \bar{\kappa}_{rs}$ for r, s both positive integers, so that we may write

$$\beta = (\bar{\kappa}_{p_1, p_2+1}) / (\bar{\kappa}_{p_1+1, p_2}), \quad \bar{\kappa}_{p_1+1, p_2} \neq 0.$$

Having expressed β in terms of the cumulants of the *observables*, a consistent estimate can readily be obtained from the sample data through moments or Fisher's k statistics. For small samples, of course, the freedom of choice for p_1, p_2 allows different estimates (the difference tending to zero in probability). The theory breaks down when the X 's are not inherently related or when they are jointly normally distributed, as the denominator is zero in either case (normality implying underidentification).

While the above exposition is concerned with a special case only and uses a different method of argument, it serves to illustrate the ingenuity and power of these early papers. Indeed, an asymptotic sampling theory was developed in the 1943c paper which also discussed a possible extension for the normal case. The method has been discussed by several leading authorities such as Kendall and Stuart (1973), Malinvaud (1966) and Madansky (1976). The method is elegant and treats variables symmetrically. Some doubt exists however about high variances in practice, presumably, for example, if the X 's were nearly normal (see Kendall and Stuart, p. 412, who also make the interesting observation that if the distribution of the x 's is symmetrical, the denominator would vanish unless $p_1 + p_2$ is chosen odd, and Malinvaud, p. 358). Attention to the nonlinear case is paid in Geary (1942b, 1949a, 1953). Geary (1963) represents Geary's views on the method some twenty years later.

Every student of statistics learns at an early stage about skewness and kurtosis and in particular of their respective measures $\sqrt{\beta_1} = \mu_3 / \mu_2^{-3/2}$ and $\beta_2 = \mu_4 / \mu_2^{-2}$ where μ_r is the r th moment about the mean of the variate in question. If the

distribution is symmetrical then $\sqrt{\beta_1}$ is zero and if the distribution is normal, then $\beta_2 = 3$. Geary has made significant contributions to the use of $\sqrt{b_1}$ and b_2 , the sample analogues of $\sqrt{\beta_1}$ and β_2 in testing for normality, a question which many of the greatest names in statistics (K. Pearson, R. A. Fisher, J. Wishart) had tackled. Writing $m_r = 1/n \sum (x - \bar{x})^r$, Geary (1933) proved the important result that \bar{x} , m_2 and $m_r m_2^{-r/2}$ are independent for normal samples.

Thus
$$E m_r^p m_2^{-rp/2} E m_2^{p/2} = E m_r^p$$
 so that $E(m_r m_2^{-r/2})^p$ can be deduced from $E m_r^p$ and $E m_2^{p/2}$ and it follows fairly readily that $E\sqrt{b_1} = 0$ and $E b_2 = 3 - 6/(n+1)$.

Expressions for the variances also follow from Geary's observation (Cramer, 1946, p. 386). Indeed, Geary and Worlledge (1947) found the seventh moment of b_2 . (The 1933 paper was in fact an important precursor of the much later Durbin-Watson theory.) By this time large sample approximations to the distributions of $\sqrt{b_1}$ and b_2 had been calculated by E. S. Pearson (see also Geary (1947a)). However, Geary (1935b), doubting their reliance for small or moderate samples, suggested that the ratio of mean deviation to standard deviation computed from the origin might be used as a test of normality and he gave 1 and 5% probability points for this statistic (subsequently known as Geary's ratio-Kendall (1956, p. 5)-for normal samples of 6-100).

Defining $a(c) = (1/n \sum |x - \bar{x}|^c) / (1/n \sum (x - \bar{x})^2)^{c/2}$,

Geary (1935c) showed that there was a high negative correlation between $a(1)$ and b_2 for normal samples and argued therefrom that $a(1)$ should be nearly as efficient as b_2 in testing for normality. Geary (1936b) gave tables for $a(1)$ and Geary and Pearson (1938) gave tables and diagrams for $a(1)$, $\sqrt{b_1}$ and b_2 . This work culminated in the technical *tour-de-force*, Geary (1947c). This paper began by considering the effects of non-normality on standard tests for differences between means and differences between variances. Theoretical results were obtained, based on truncated expansions, which indicated the kind of departures from normality likely to have serious effects on the tests. Regarding the *t*-test, for example, "the standard table probabilities can be seriously at variance with the true probabilities when the universes from which the samples are drawn are markedly asymmetrical" (p. 217). It was then established that the expression $a(c)$ above is asymptotically normal provided the r th absolute moment about the origin is finite where $r = \max(2c, 4)$. Close approximation to the first four raw moments of $a(c)$ for normal populations were deduced and $a(c)$ carefully considered as a test for kurtosis for varying c for both small and large samples. It turned out that $a(c)$ is a useful test for a broad range of c although $a(4) = b_2$ is most efficient for large samples relative to a large class of alternative populations (a fascinating result).

Tests of asymmetry were also suggested using

$$g(c) = 1/n \left\{ \sum_{x \geq \bar{x}} (x - \bar{x})^c - \sum_{x < \bar{x}} |x - \bar{x}|^c \right\} / \left\{ 1/n \sum (x - \bar{x})^2 \right\}^{c/2}$$

The first two raw moments of $g(c)$ were calculated for normal populations, and for a given field of alternative populations it was demonstrated that $g(3) = \sqrt{b_1}$ is the most efficient test for large samples although $g(c)$, $2 \leq c \leq 5$, are almost as good.

The paper ends with a plea to applied workers to beware of assuming normality uncritically, and a plea to other theorists to extend the results to wider classes of alternative populations. The work illustrates excellently Geary's flair and sensitivity for statistical problems and also his enormous energy and technical ability. Aitchison and Brown (1957) briefly discuss the use of the Geary tests as tests for lognormality.

Two other papers predating the 1947c paper must be mentioned as both developed beautiful results which are now regarded as classics. Geary (1936a), in a consideration of the robustness of the t -test which anticipated some results of (1947c) and which involved expansions to terms in n^{-2} of the first four moments of t also showed that if \bar{x} and $\Sigma(x - \bar{x})^2$ are independent for populations possessing finite cumulants of all orders, then x must be normally distributed. This was a deep converse to what was well known and in fact was subsequently generalised by Lukacs (1942) under less restrictive conditions on the existence of cumulants. Later, Geary (1942a) proved a property of maximum likelihood estimation which has become equally well-known, that is, under regularity conditions the maximum likelihood estimators minimise the generalised variance for large samples (an extension of the similar earlier result of Fisher for the single parameter case). Furthermore a theorem on the efficiency of sufficient statistics was also demonstrated. Let $\hat{\theta}$ and θ^* be unbiased normally distributed estimators of the $k \times 1$ vector θ . Then if $\hat{\theta}$ is sufficient with respect to θ while θ^* is not sufficient, the generalised variance of the $\hat{\theta}$ is less than that of the θ^* (again a generalisation of an earlier Fisher result).

Geary (1948), following a suggestion of Richard Stone, compared proposed estimators of Tinbergen and Koopmans with Hotelling's principal components theory and went on to develop an estimator more efficient than a broad set of alternatives under certain conditions. This paper involved some work on Reiersöl's instrumental variables introduced in relation to confluence analysis (Reiersöl (1941) and (1945)). This led to the famous Geary (1949a) paper on instrumental variable (IV) estimation in the context of errors in variables, a paper which has led to Geary being considered as the leading pioneer, with Reiersöl of the IV method (see Malinvaud (1966), p. 347, Brundy and Jorgenson (1971), p. 207). This is possibly Geary's most quoted paper and it begins by considering the problem of estimating $\beta = -\beta_1/\beta_2$ in the model $\beta_1 X_1 + \beta_2 X_2 = 0$ where the X 's are observed with error, the observables x_i being equal to $X_i + u_i$, as in the method of cumulants described above. (The case, more often treated today, where the relationship itself contains an error term was definitively analysed in Sargan (1958) for the single equation case.) Let z be the IV measured without error and suppose X_1, X_2, z have zero means, are joint normal and temporally uncorrelated. Geary writes the IV estimator of β as

$$b = (\Sigma x_2 z) / (\Sigma x_1 z) \quad (\equiv \text{num/den, say})$$

and shows, using Geary (1944b), that its density can be written

$$\Phi(b) db = \frac{\{(n-2)/2\}! (1+y^2)^{-n/2}}{\{(n+3)/2\}! \sqrt{\pi}} dy$$

where

$$y = \left\{ \frac{Ez^2(b^2Ex_1^2 - 2bEx_1x_2 + Ex_2^2)}{(bEx_1z - Ex_2z)^2} - 1 \right\}^{-\frac{1}{2}}$$

so that $y\sqrt{n-1}$ is distributed as t with $n-1$ degrees of freedom. By considering confidence intervals, it is shown that the precision of the method (asymptotically) is improved if z is chosen as highly correlated with X_1 and X_2 as possible and the idea of finding an optimal combination of instruments is treated.

The theory is then extended to the time series case where the observables are taken as non-stochastic with the errors normally distributed, independent of each other and of the other variables. The estimator b in this case is the ratio of two independent normal variables so that Geary (1930b) applies by which $\{bE(\text{den}) - E(\text{num})\} / \sqrt{\{b^2 \text{var}(\text{den}) + \text{var}(\text{num})\}}$ is a $N(0, 1)$ variate. Of course, this expression involves unknown error variances while the previous expression only involves expectations of functions of observables which can easily be consistently estimated. Although the first model would seem to be inapplicable to the time series "sequences of n " case it is found that the operative first theory can be used with confidence in such a case for moderately sized samples with error variances not too large. This brilliant paper ends with further asymptotic theory, an application of the first model to some US data and interesting miscellaneous comments.

Geary's work continued through the 1950s and 1960s with undiminished vigour. Thus Geary (1950-51) discovered the utility function from which the Stone system of demand functions would follow (the utility function is nowadays known as the Stone-Geary function). Geary (1958) criticised the use of official exchange rates for making international comparison in real terms of national output, the latter being an aggregate of many economic flows of different types. Appropriate conversion factors, one for each country, were derived from a set of linear homogenous equations. Interestingly, in the two country case, the ratio of conversion factors was a simple weighted price index, where the weights were harmonic means of the quantities. Geary (1961a) devised a method for estimating the realised trade gain (g) in goods and services from changes in terms of trade expressed at base year prices. It was proposed to measure g by the formulae

$$\begin{aligned} g &= M(1/p(m) - 1/p(x)) \text{ for } X > M \\ &= X(1/p(m) - 1/p(x)) \text{ for } X < M \end{aligned}$$

where X, M are exports and imports respectively at current prices while $p(x)$ and $p(m)$ are the corresponding deflators (see Hibbert, 1975, for discussion of the method).

Geary (1966a) introduced the concept of average critical value (ACV) as a means of ranking tests concerning a parameter θ in order of sensitivity. Using a statistic S to test $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ the sensitivity of S is high (it is highly efficient) if equating $E(S | \theta_1)$ to the boundary point of the critical region of the test yields a value of θ_1 (the ACV) which is close to θ_0 . This notion is compared with the classical power function approach which, of course, requires calculation

of the distribution of S , given H_1 , and examples are considered. Later, Stuart (1967) showed under fairly general conditions that, given two test functions, the ratio of the ACV's for these tests is an approximation to a power, commonly the square root, of their asymptotic relative efficiency.

Geary and Leser (1968) demonstrated that in regression, all the coefficients could be significant with R^2 insignificant (and vice versa). Geary (1969a) in a study of forecasting showed that successful estimation in the sense of good R^2 and DW statistic did not imply that the equation would be better for forecasting than a less successful equation (an empirical demonstration) while Geary (1969c) compared the efficiency of maximum likelihood and *ex ante* reduced form (estimating the reduced form (RF) directly using the structure merely to deduce the position of zeroes in the RF matrix) for forecasting in the context of a simple two equation recursive model. Geary (1970) proposed a simple alternative to the DW test based on a count of sign changes in the series of calculated residuals from a regression. Geary (1971) used order statistics theory to find a test for identifying outliers in regression analysis and, finally, Geary (1972d) is a reasoned plea to theorists to pay rather more attention to efficiency relative to consistency.

While hoping that this survey has presented an adequate picture of Geary's main theoretical papers, I am sadly aware that I cannot do him justice as an individual. He commands a beautiful, flowing prose style and has a ready sense of humour. Although a confrère of men like Sir Ronald Fisher and Ragnar Frisch, yet he listens as carefully and courteously to the seminar contribution of the neophyte as to that of the leading expert. Finally, however, I would return to his work, with its rare blend of high theory, common sense and feel for real problems. Even his most abstract papers contain applications and examples. He has in abundant measure the vital qualities of curiosity and imagination. Long may they continue.

New University of Ulster.

ACKNOWLEDGEMENTS

The author wishes to thank F. P. Murphy for locating some of the papers and Susan Thom for some assistance with stylistic detail.

REFERENCES

Note: A complete list of publications by R. C. Geary is published in this Review. Articles referred to by Professor Spencer in his evaluation of Geary's scientific writings are marked with an asterisk in that list.

AITCHISON, J. and J. A. C. BROWN, 1957, *The Lognormal Distribution*, Cambridge: Cambridge University Press.

- BRUNDY, J. M. and D. W. JORGENSON, 1971, "Efficient Estimation of Simultaneous Equations by Instrumental Variables", *The Review of Economics and Statistics*, Vol. LIII, No. 3, August.
- CRAMER, H., 1937, *Random Variables and Probability Distributions*, Cambridge: Cambridge University Press.
- CRAMER, H., 1946, *Mathematical Methods of Statistics*, Princeton: Princeton University Press.
- HIBBERT, J., 1975, "Measuring Changes in the Nation's Real Income", *Economic Trends*, No. 255, January.
- KENDALL, M. G., 1956, *Exercises in Theoretical Statistics*, Second Edition, London: Charles Griffin and Co., Ltd.
- KENDALL, M. G. and A. STUART, 1973, *The Advanced Theory of Statistics*, Volume 2, Third Edition, London: Charles Griffin and Co., Ltd.
- LUKACS, E., 1942, "A Characterization of the Normal Distribution", *Annals of Mathematical Statistics*, Vol. 13.
- MADANSKY, A., 1976, *Foundations of Econometrics*, Amsterdam: North Holland.
- MALINVAUD, E., 1966, *Statistical Methods of Econometrics*, Amsterdam: North Holland.
- MARSAGLIA, G., 1965, "Ratios of Normal Variables and Ratios of Sums of Uniform Variables", *Journal of the American Statistical Association*, Vol. 60.
- REIERSÖL, O., 1941, "Confluence Analysis by Means of Lag Moments and Other Methods of Confluence Analysis", *Econometrica*, Vol. 9, No. 1.
- REIERSÖL, O., 1945, "Confluence Analysis by Means of Sets of Instrumental Variables", *Arkiv for Matematik, Astronomi och Fysik*, Vol. 32.
- SARGAN, J. D., 1958, "The Estimation of Economic Relationships using Instrumental Variables", *Econometrica*, Vol. 26, July.
- SCADDING, J. L., 1973, "The Sampling Distribution of the Liviatan Estimator of the Geometric Distributed Lag Parameter", *Econometrica*, Vol. 41, No. 3, May.
- STUART, A., 1967, "The Average Critical Value Method and the Asymptotic Relative Efficiency of Tests", *Biometrika*, Vol. 54, Nos. 3 and 4.