DOCUMENT RESUME

ED 055 120 TM 000 858

AUTHOR Kirk, David B.

TITLE A Method for Approximating the Bivariate Normal Correlation Coefficient.

PUB DATE 1 Aug 71

NOTE 5p.; Supplement to RB-71-35

EDRS PRICE MF-$0.65 HC-$3.29

DESCRIPTORS *Algorithms; *Computer Programs; *Correlation; Mathematical Models; *Mathematics; Statistical Analysis

IDENTIFIERS Gaussian Quadrature; *Newton Raphson Iteration Technique

ABSTRACT

Improvements of the Gaussian quadrature in conjunction with the Newton-Raphson iteration technique (TM 000 789) are discussed as effective methods of calculating the bivariate normal correlation coefficient. (CK)

ERIC

ED 055120

A Method for Approximating the Bivariate Normal Correlation Coefficient

Supplementary Paper
to RB-71-35

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

David B. Kirk

August 1, 1971

TM 000 858

ERIC

It has been shown in the RB-71-35 that Gaussian quadrature supplemented by a Newton-Raphson iteration technique is an effective method of calculation of the Bivariate Normal (Tetrachoric) r . Subsequent studies made it apparent that significant improvements could easily be incorporated into the calculation with relatively little increase in complexity, or cost, of the computational technique. These improvements are discussed in this supplement.

1. Estimates of the standard deviates h and k

The value of k may be written in the form

$$\Phi\ (k) = \frac{1}{\sqrt{2\pi}} \int_{o}^{k} e^{-\frac{x^2}{2}}\ dx$$

where, in the notation of the original paper, Φ (k) is equivalent to .5 minus the marginal percentage, q_1. In the original model, Hastings' approximation without any modification was used to estimate k, although a remark was made that the result could be improved by an iteration technique.

It is apparent from this form of the integral that we are faced with precisely the same problem as in our evaluation of r , namely, we must compute (or estimate) a variable upper limit of a definite integral. There is no reason, therefore, not to use the same algorithmic technique, i.e., Gaussian quadrature and Newton-Raphson iteration, to improve Hastings' estimates. This seems even more feasible when one realizes the necessary calculation ingredients, the Gaussian quadrature coefficients, the related weights, and the iteration structure are already available for the evaluation of r .

2. The Gaussian Quadrature.

The 5-point quadrature used in the original study was quite rapid and gave acceptable values except where the joint and marginal values were close. However, an increase of only 3 points to an 8-point quadrature resulted in the convergence of many values which previously had failed. This increased accuracy is important not only in the evaluation of the final r integral, but of equal and perhaps greater benefit in establishing more accurate values of the h and k parameters which make up the function. Consider the following table:

Different Quadrature Effects on h and k Calculation

Area	h true	Hastings' Estimate Unmodified	5-Point Quadrature		8-Point Quadrature	
			Value	Iterations	Value	Iterations
.5	0	$-1.01\cdot 10^{-7}$	$-.3\cdot 10^{-13}$	1	$-.4\cdot 10^{-13}$	1
.158655254	1	.999968	1.0000004	2	1.0000002	2
.022750132	2	2.000435	2.000002	2	2.000001	2
.001349898	3	3.000314	3.00022	2	2.999990	2

Thus, a substantial improvement in the values of h and k is achieved with only two iterations.

3. The Starting Estimate

In accordance with the above improvements, two terms of the series expansion were used instead of one, and the resulting quadratic equation in r solved to provide a better starting estimate. Extreme values again caused this estimate to exceed 1, consequently it was necessary to set limiting values as was done previously. Again, no one value seemed to assure convergence over the entire range of r. For example, a P value of .001131 (h = 2, k = 3, r = .80) failed to converge with a starting estimate of .97 but converged to

.8003 readily with a lower value. On the other hand, a P value of .477473 ($h = 0$, $k = 0$, $r = .99$) failed with a starting estimate of .90 but converged in 5 iterations to .99096 with a starting estimate of .97. On the assumption that the majority of r calculations will be within the range $-.80 < r < .80$ and only occasionally near the extreme values which tend to give the most computational difficulty, the bounds were set at $\pm .80$ with a final pass using $\pm .97$ if the first fails to converge. Fairly extensive testing has resulted in the convergence of all "reasonable" values by this method.

4. The Convergence Criteria

Two convergence values were used in the attached examples: $1 \cdot 10^{-5}$ for both h and k calculations and $1 \cdot 10^{-4}$ for the r calculation. The effect of these values is, of course, evident in the above table.

Summary

Three versions of the algorithm are thus readily available for use:

1) 5-point quadrature, unmodified Hastings' estimates of h and k.
2) 5-point quadrature, improved estimates of h and k.
3) 8-point quadrature, improved estimates of h and k.

The attached sheet of computer output indicates the range of values for the 8-point quadrature. An average calculation for the 8-point quadrature required .027 seconds per computed value of r vs. .020 for a 5-point quadrature. On 50 x 50 matrix of "live" data an average r required .018 seconds using 8-point quadrature compared with .016 seconds for 5 points. The stability of the higher quadrature seems to justify its use.

P	Q1	Q2	R CALC	R TRUE
0.11375000D-01	0.500000000D 00	0.227500000D-01	0.0	0.0
0.79328000D-01	0.500000000D 00	0.158655254D 00	0.38640D-05	0.0
0.67500000D-03	0.500000000D 00	0.135000000D-02	0.0	0.0
0.13518000D-01	0.227501320D-01	0.500000000D 00	0.99991D-01	0.10
0.50000000D-05	0.134989800D-02	0.134989800D-02	0.10205D 00	0.10
0.84900000D-03	0.134989800D-02	0.500000000D 00	0.99758D-01	0.10
0.88981000D-01	0.158655254D 00	0.500000000D 00	0.10000D 00	0.10
0.31549500D 00	0.500000000D 00	0.500000000D 00	0.40000D 00	0.40
0.12739800D 00	0.158655254D 00	0.500000000D 00	0.50000D 00	0.50
0.10370000D-02	0.158655254D 00	0.134989800D-02	0.50024D 00	0.50
0.46000000D-03	0.227500000D-01	0.135000000D-02	0.49988D 00	0.50
0.33333300D 00	0.500000000D 00	0.500000000D 00	0.50000D 00	0.50
0.13490000D-02	0.134989800D-02	0.500000000D 00	0.70756D 00	0.70
0.39758400D 00	0.500000000D 00	0.500000000D 00	0.80000D 00	0.80
0.15309100D 00	0.158655254D 00	0.500000000D 00	0.80001D 00	0.80
0.13490000D-02	0.158655254D 00	0.134989800D-02	0.84872D 00	0.85
0.41169900D 00	0.500000000D 00	0.500000000D 00	0.85000D 00	0.85
0.22749000D-01	0.227501320D-01	0.500000000D 00	0.87145D 00	0.90
0.15794900D 00	0.158655254D 00	0.500000000D 00	0.89996D 00	0.90
0.42821700D 00	0.500000000D 00	0.500000000D 00	0.90000D 00	0.90
0.22742000D-01	0.227501320D-01	0.158655254D 00	0.95016D 00	0.95
0.22750000D-01	0.227501320D-01	0.500000000D 00	0.89829D 00	0.95
0.44945900D 00	0.500000000D 00	0.500000000D 00	0.95004D 00	0.95
0.15863100D 00	0.500000000D 00	0.158655000D 00	0.94967D 00	[illegible]
0.12813000D 00	0.158655254D 00	0.158655254D 00	0.95004D [illegible]	[illegible]95
0.16024000D-01	0.227501320D-01	0.227501320D-01	0.95003D 00	0.95
0.13490000D-02	0.227501320D-01	0.134989800D-02	0.95138D 00	0.95
0.80900000D-03	0.134989800D-02	0.134989800D-02	0.95001D 00	0.95
0.15863100D 00	0.500000000D 00	0.158655254D 00	0.94960D 00	0.95
0.15865500D 00	0.158655254D 00	0.500000000D 00	0.95988D 00	0.97
0.22750000D-01	0.158655254D 00	0.227501320D-01	0.95996D 00	0.97
0.47747300D 00	0.500000000D 00	0.500000000D 00	0.99096D 00	0.99
0.11020000D-02	0.134989800D-02	0.134989800D-02	0.99105D 00	0.99
0.19712000D-01	0.227501320D-01	0.227501320D-01	0.99098D 00	0.99
0.14500300D 00	0.158655254D 00	0.158655254D 00	0.99097D 00	0.99
0.24203890D 00	0.500000000D 00	0.500000000D 00	-0.50000D-01	-0.05
0.46000000D-05	0.227501320D-01	0.134989800D-02	-0.19949D 00	-0.20
0.11790000D-03	0.227501320D-01	0.500000000D 00	-0.75001D 00	-0.75
0.20000000D-06	0.134989800D-02	0.500000000D 00	-0.74986D 00	-0.75
0.91563000D-02	0.158655254D 00	0.500000000D 00	-0.75000D 00	-0.75
0.11502670D 00	0.500000000D 00	0.500000000D 00	-0.75000D 00	-0.75
0.70480000D-03	0.500000000D 00	0.158655254D 00	-0.90001D 00	-0.90
0.10000000D-06	0.227501320D-01	0.500000000D 00	-0.90165D 00	-0.90
0.71783100D-01	0.500000000D 00	0.500000000D 00	-0.90000D 00	-0.90
0.50541300D-01	0.500000000D 00	0.500000000D 00	-0.95004D 00	-0.95
0.45167200D-01	0.500000000D 00	0.500000000D 00	-0.96008D 00	-0.96
0.51000000D-05	0.158655254D 00	0.500000000D 00	-0.95711D 00	-0.96
0.40000000D-06	0.158655254D 00	0.500000000D 00	-0.95985D 00	-0.97
0.39083000D-01	0.500000000D 00	0.500000000D 00	-0.97016D 00	-0.97
0.22526700D-01	0.500000000D 00	0.500000000D 00	-0.99096D 00	-0.99

ELAPSED TIME IN MICROSECONDS 0.1381120 07