APPROXIMATIONS TO MULTIVARIATE NORMAL ORTHANT PROBABILITIES[1]

By Ralph Hoyt Bacon

GPL Division—General Precision, Inc.[2]

1. Summary. In this work, a close approximation to the positive normal orthant is obtained for the special case in which all the correlation coefficients are equal. From this expression, an approximation for the general case is suggested. For special cases, even closer approximations are obtainable.

2. Statement of the problem. Consider n correlated random variables, distributed with zero means and unit variance according to a multivariate normal distribution, whose density function we will write as $\phi_n(x_1, x_2, \cdots x_n)$. Then the probability that all n random variables will be positive is

$$\int_0^\infty \int_0^\infty \cdots \int_0^\infty \phi_n(x_1, x_2, \cdots x_n)\, dx_1\, dx_2 \cdots dx_n .$$

This probability will be denoted by $P_n(c_{11}, c_{12}, \cdots c_{nn})$, where c_{11}, etc., denote the elements of the inverse of the covariance matrix. Let the correlation coefficients be denoted by ρ_{ij}. Then, since the density function obeys the relation

$$\frac{\partial \phi_n}{\partial \rho_{ij}} = \frac{\partial^2 \phi_n}{\partial x_i\, \partial x_j},$$

we have, for example,

$$\begin{aligned} \frac{\partial P_n}{\partial \rho_{12}} &= \int_0^\infty \int_0^\infty \cdots \int_0^\infty \left[\int_0^\infty \int_0^\infty \frac{\partial^2 \phi_n}{\partial x_1\, \partial x_2}\, dx_1\, dx_2 \right] dx_3 \cdots dx_n \\ &= \frac{1}{2\pi(1 - \rho_{12}^2)^{\frac{1}{2}}} P_{n-2}(c_{33}, c_{34}, \cdots c_{nn}). \end{aligned} \tag{1}$$

There will be $\frac{1}{2}n(n-1)$ such equations—one for each correlation coefficient. Equation (1) is a special case of Equation (6) given by Plackett [5].

3. Equal correlation coefficients. Let the resultant probability in the special case in which all the correlation coefficients are equal to be denoted by $P_n(\rho)$. Then the $\frac{1}{2}n\ (n-1)$ Equations (1) add up to

$$\frac{dP_n(\rho)}{d\rho} = \frac{n(n-1)}{4\pi(1-\rho^2)^{\frac{1}{2}}} P_{n-2}\left(\frac{\rho}{1+2\rho}\right).$$

Therefore, since $P_n(0) = (\frac{1}{2})^n$, we have

$$P_n(\rho) = \left(\frac{1}{2}\right)^n + \frac{n(n-1)}{4\pi} \int_0^\rho P_{n-2}\left(\frac{\lambda}{1+2\lambda}\right) \frac{d\lambda}{(1-\lambda^2)^{\frac{1}{2}}}. \tag{2}$$

Equation (2) is the same as Equation (102′) given by Ruben [6].

Received May 25, 1961; revised September 11, 1962.

[1] Some of these results were presented at the Atlantic City Meeting, September, 1957 [2].

[2] Formerly General Precision Laboratory.

Now, P_1, P_2, and P_3 can be found by direct integration; the results are

$$P_1 = \tfrac{1}{2}, \tag{3}$$

$$P_2 = \tfrac{1}{4}\,[1 + (2/\pi)\text{ arc sin } \rho], \tag{4}$$

$$P_3 = \tfrac{1}{8}\,[1 + (2/\pi)\ (\text{arc sin } \rho_{12} + \text{arc sin } \rho_{23} + \text{arc sin } \rho_{13})]. \tag{5}$$

If we combine (2) with (4), we obtain

$$P_4(\rho) = \frac{1}{16}\left[1 + \frac{12}{\pi}\text{ arc sin } \rho + \frac{24}{\pi^2}\int_0^\rho \text{arc sin } \frac{\lambda}{1 + 2\lambda}\,\frac{d\lambda}{(1 - \lambda^2)^{\frac{1}{2}}}\right], \tag{6}$$

and if we combine (2) with (5) (remembering that we are considering only the case for which $\rho_{12} = \rho_{13} = \rho_{23} = \rho$) we obtain

$$P_5(\rho) = \frac{1}{32}\left[1 + \frac{20}{\pi}\text{ arc sin } \rho + \frac{120}{\pi^2}\int_0^\rho \text{arc sin } \frac{\lambda}{1 + 2\lambda}\,\frac{d\lambda}{(1 - \lambda^2)^{\frac{1}{2}}}\right]. \tag{7}$$

We can continue the process of combining (2) with the succeeding equations, and obtain the general equation

$$\begin{aligned} P_n(\rho) = \left(\frac{1}{2}\right)^n \Bigg[1 &+ \frac{n(n-1)}{\pi}\text{ arc sin } \rho \\ &+ \frac{n(n-1)(n-2)(n-3)}{\pi^2}\int_0^\rho \text{arc sin } \frac{\lambda}{1 + 2\lambda}\,\frac{d\lambda}{(1 - \lambda^2)^{\frac{1}{2}}} \\ &+ \frac{n(n-1)\cdots(n-5)}{\pi^3}\int_0^\rho\int_0^{\frac{\mu}{1+2\mu}} \text{arc sin } \frac{\lambda}{1 + 2\lambda}\cdot\frac{d\lambda}{(1 - \lambda^2)^{\frac{1}{2}}}\,\frac{d\mu}{(1 - \mu^2)^{\frac{1}{2}}} \\ &+ \frac{n(n-1)\cdots(n-7)}{\pi^4}\int_0^\rho\int_0^{\frac{\nu}{1+2\nu}}\int_0^{\frac{\mu}{1+2\mu}} \text{arc sin } \frac{\lambda}{1 + 2\lambda} \\ &\cdot\frac{d\lambda}{(1 - \lambda^2)^{\frac{1}{2}}}\,\frac{d\mu}{(1 - \mu^2)^{\frac{1}{2}}}\,\frac{d\nu}{(1 - \nu^2)^{\frac{1}{2}}} + \cdots\Bigg]. \end{aligned} \tag{8}$$

Equation (8) is an exact expression for $P_n(\rho)$. The integrals included in this expression cannot, as far as I am aware, be expressed in terms of elementary functions, but they can be found by numerical integration, and a short table of them is shown in Table I.

For brevity, let us denote these integrals by $I_2(\rho)$, $I_3(\rho)$, etc. [$I_1(\rho)$, in this notation, is just simply arc sin ρ]; and let us denote our working approximations to them by I_2^*, I_3^*, etc. Also, for brevity, let $\theta = \text{arc sin } \rho/\pi$. Now, $I_2(0) = 0$; $I_2(\rho)$ is nearly equal to $(\text{arc sin } \rho)^2/2$ when ρ is very small;

$$I_2(\tfrac{1}{2}) = \pi^2/120;$$

$$I_2(1) = \pi^2/24.$$

If we now let

$$I_2^*(\rho) = [(\text{arc sin } \rho)^2/2(1 + 4\theta)], \tag{9}$$

TABLE I

The Integrals Appearing in Equation (8)

ρ	$I_2(\rho)$	$I_3(\rho)$	$I_4(\rho)$
.00	.000000	.000000	.000000
.05	.001172	.000017	.000002
.10	.004404	.000117	.000022
.15	.009477	.000343	.000083
.20	.016067	.000712	.000203
.25	.024057	.001232	.000397
.30	.033375	.001907	.000673
.35	.043812	.002727	.001037
.40	.055459	.003706	.001495
.45	.068254	.004843	.002052
.50	.082247	.006152	.002714
.55	.097454	.007628	.003486
.60	.114012	.009291	.004379
.65	.132053	.011154	.005404
.70	.151813	.013243	.006580
.75	.173640	.015607	.007934
.80	.198120	.018306	.009506
.85	.226180	.021464	.011370
.90	.259820	.025310	.013668
.95	.303950	.030429	.016770
1.00	.411234	.043064	.024159

we find that the expression is exact when $\rho = 0, \frac{1}{2}$, or 1. The error of the approximation is shown in Fig. 1.

Similarly, $I_3(0) = 0$; $I_3(\rho)$ is nearly equal to $(\text{arc sin } \rho)^3/6$ when ρ is very small;

$$I_3(\tfrac{1}{2}) = \pi^3/5040;$$

$$I_3(1) = \pi^3/720.$$

Therefore, the expression

$$I_3^* = [(\text{arc sin } \rho)^3/6(1 + 4\theta)(1 + 8\theta)] \tag{10}$$

is exact when $\rho = 0, \frac{1}{2}$, or 1. A graph of the error of this approximation would be similar in shape to the curve in Fig. 1.

In general, $I_K(0) = 0$; $I_K(\rho)$ is nearly equal to $(\text{arc sin } \rho)K/K!$ when ρ is very small;

$$I_K(\tfrac{1}{2}) = \pi^K/(2K + 1)!;$$

$$I_K(1) = \pi^K/(2K)!,$$

and the expression

$$I_K^*(\rho) = \frac{(\text{arc sin } \rho)^K}{K!(1 + 4\theta)(1 + 8\theta) \cdots [1 + 4(K - 1)\theta]} \tag{11}$$

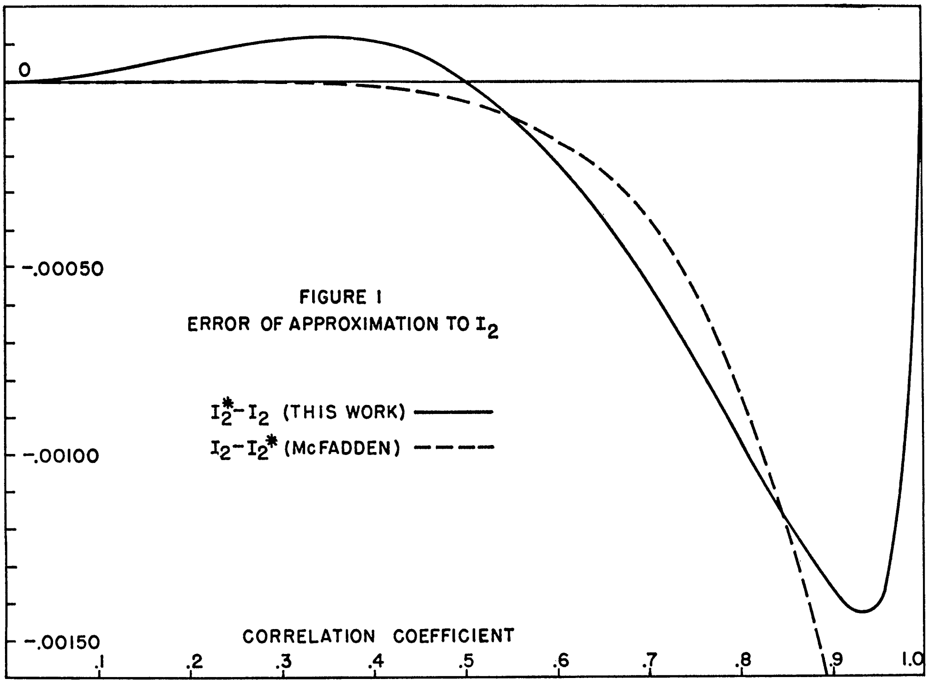

FIG. 1.

is exact when $\rho = 0, \frac{1}{2}$, or 1, and provides a convenient approximation to $I_K(\rho)$ as ρ varies from 0 to 1.

Equation (8) is now replaced by the approximation

$$P_n(\rho) \approx \left(\frac{1}{2}\right)^n \left[1 + \frac{n(n-1)}{\pi} \text{arc sin } \rho + \frac{n(n-1)(n-2)(n-3)}{\pi^2} I_2^* + \frac{n(n-1)\cdots(n-5)}{\pi^3} I_3^* + \cdots\right]. \tag{12}$$

4. The approximation suggested by (12) for the general case. An approximation for the general case is suggested by the coefficients of the various terms in in (12). For

$$n(n-1)/2$$

is the number of correlation coefficients; $[n(n-1)(n-2)(n-3)]/2^2\,2!$ is the number of disjoint pairs of correlation coefficients—that is, the number of pairs of the form $\rho_{ab}\,\rho_{cd}$ but excluding pairs of the form $\rho_{ab}\,\rho_{bc}$;

$$[n(n-1)\cdots(n-5)]/2^2\,3!$$

is the number of disjoint triplets;

$$[n(n-1) \cdots (n-7)]/2^4 4!$$

is the number of disjoint quartets; and so on.

Given a pair of correlation coefficients, ρ_{ab} and ρ_{cd}, the number of distinct additional correlation coefficients obtained by interchanging these four subscripts is four: ρ_{ac}, ρ_{ad}, ρ_{bc}, and ρ_{bd}; the number of distinct additional coefficients obtained by interchanging the six subscripts in a given disjoint triplet is 12; by interchanging the 8 subscripts in a given disjoint quartet is 24; and so on. If we expand the denominators of the successive terms in (12), the coefficients of θ in the resultant polynomials are 4, 12, 24, etc. It would therefore appear that the following expression is a reasonable approximation for the general case:

$$P_n \approx \left(\frac{1}{2}\right)^n \Bigg[1 + 2\sum \theta + 4 \sum \frac{\theta_{ab}\theta_{cd}}{1 + \sum \theta_{ij}} + 8 \sum \frac{\theta_{ab}\theta_{cd}\theta_{ef}}{(1 + \frac{1}{3}\sum \theta_{ij})(1 + \frac{2}{3}\sum \theta_{ij})} + \cdots$$

$$(13) \qquad + 2^K \sum \frac{\theta_{ab}\theta_{cd} \cdots \theta_{gh}}{\left(1 + \frac{2\sum \theta_{ij}}{K(K-1)}\right)\left(1 + \frac{4\sum \theta_{ij}}{K(K-1)}\right) \cdots \left(1 + \frac{2\sum \theta_{ij}}{K}\right)} + \cdots \Bigg],$$

where, in each case, $\sum\theta_{ij}$ means the sum obtained by interchanging the subscripts in the numerator—that is, $\sum\theta_{ij}$ does not contain the θ's that are in the numerator.

Equation (13) is a working approximation obtained by very non-rigorous methods. In the next section we compare it with some exact results; and in the following sections, we compare it with other approximations.

5. Comparison of (13) **with some exact figures.** It follows from the work of Anis and Lloyd [1] that, when the inverse of the covariance matrix has 2's on the main diagonal, -1's on the two adjacent diagonals, and zeros elsewhere, then $P_n = 1/(1+n)$. The comparison between (13) and a few of these exact probabilities follows:

n	P_n exact	P_n Equation (13)
4	1/5	.199818
5	1/6	.166165
6	1/7	.142577

Thus, the approximation is, in this special case, good enough for many useful purposes.

TABLE II

Comparison of the Three Proposed Approximations with Some Exact Figures

ρ_{12}	ρ_{23}	ρ_{34}	ρ_{14}	ρ_{13}	ρ_{24}	P_4			
						Exact	Eq. (13)	Plackett's Approx.	Eq. (14)
.926	.135	.809	0	0	0	.181502	.179507	.180672	.181943
.809	.309	.809	0	0	0	.175417	.170545	.172500	.175000
$(5/8)^{\frac{1}{2}}$	1/4	$(5/8)^{\frac{1}{2}}$	0	0	0	.167675	.164602	.166171	.168014
$(1/2)^{\frac{1}{2}}$	1/2	$(1/2)^{\frac{1}{2}}$	0	0	0	.166667	.159226	.161458	.164583
1/2	.309	.926	0	0	0	.160997	.157150	.158575	.160318
1/2	1/2	.809	0	0	0	.158407	.152381	.154167	.156667
1/2	1/2	$(5/8)^{\frac{1}{2}}$	0	0	0	.156295.	.150808	.152538	.154954
1/2	1/2	$(1/2)^{\frac{1}{2}}$	0	0	0	.148438	.144345	.145833	.147917
1/2	.309	.809	0	0	0	.146944	.144697	.145833	.147222
1/2	$(1/2)^{\frac{1}{2}}$	1/2	0	0	0	.146701	.140792	.142361	.144676
.809	1/2	.309	0	0	0	.143055	.139762	.140833	.146333
1/2	.809	.309	0	0	0	.141944	.136585	.137500	.139286
1/2	$(5/8)^{\frac{1}{2}}$	1/4	0	0	0	.135973	.132270	.133025	.134396
1/2	1/2	1/2	0	0	0	.133333	.130952	.131944	.133333
.309	.926	.135	0	0	0	.130808	.128221	.128516	.129165
1/2	.809	.135	0	0	0	.129733	.127796	.128227	.128979
.309	.809	.309	0	0	0	.129583	.126923	.127500	.128571
1/2	1/2	.309	0	0	0	.121597	.120238	.120833	.121667
1/2	1/2	1/2	1/2	0	0	.166667	.156250	.166667	not relevant
1/2	1/2	1/2	0	1/2	0	.150000	.151041	.152777	not relevant
$(3/8)^{\frac{1}{2}}$	2/3	$(3/8)^{\frac{1}{2}}$	1/4	$(1/6)^{\frac{1}{2}}$	$(1/6)^{\frac{1}{2}}$	.200000	.199818	.207653	not relevant

$.926 = (3 + 5^{\frac{1}{2}})/(4 \cdot 2^{\frac{1}{2}})$ $.809 = (5^{\frac{1}{2}} + 1)/4$

$.135 = (3 - 5^{\frac{1}{2}})/(4 \cdot 2^{\frac{1}{2}})$ $.309 = (5^{\frac{1}{2}} - 1)/4$

The exact values of P_4 for a few discrete examples in which $\rho_{13} = \rho_{14} = \rho_{24} = 0$, were found by Schläfli [7], and for a few others by Coxeter [3]. The decimal equivalents of these exact fractions are listed, in order of decreasing P_4, in the first 18 lines of Table II; the next two lines are examples given by Plackett [5]; and the last line is the example that follows from the work of Anis and Lloyd [1].

The error of the approximation yielded by (13) never exceeds 0.06 P_4.

6. Comparison of (9) with approximations given by McFadden and by Sondhi. In 1956, McFadden [4] proposed the approximation

$$I_2^* = \frac{(\text{arc sin } \rho)^2(3 + 5 \text{ arc sin } \rho)}{6(1 + \text{arc sin } \rho)(1 + 2 \text{ arc sin } \rho)}.$$

The error of this approximation, for correlation coefficients smaller than .9 is shown in Fig. 1. When $\rho = 1$, the error is .00792. When ρ is smaller than $\frac{1}{2}$, the approximation is much closer than ours.

In 1961, Sondhi [8] proposed an approximation of considerable complexity yielding very great accuracy when the correlation coefficient is greater than $\frac{1}{2}$. It thus supplements McFadden's approximation.

From the point of view of our work, however, the chief feature of these two approximations is not their relative accuracy, nor complexity, but the fact that they do not (as McFadden points out) lead to approximations for the more general case.

7. Comparison of (13) **with an approximation given by Plackett** [5]. When $n = 4$, (1) becomes

$$\frac{\partial P_4}{\partial \rho_{12}} = \frac{1}{8\pi(1 - \rho_{12}^2)^{\frac{1}{2}}} \left[1 + \frac{2}{\pi} \text{arc sin } \frac{c_{34}}{(c_{33}\, c_{44})^{\frac{1}{2}}} \right].$$

There are six such equations. Upon combining them we obtain

$$P_4 = P_4(0) + \frac{1}{8\pi} \sum \int_0^{\rho_{ab}} \left[1 + \frac{2}{\pi} \text{arc sin } \frac{c_{cd}}{(c_{cc}\, c_{dd})^{\frac{1}{2}}} \right] \frac{d\lambda_{ab}}{(1 - \lambda_{ab}^2)^{\frac{1}{2}}}$$

$$= \frac{1}{16} \left[1 + \frac{2}{\pi} \sum \text{arc sin } \rho + \frac{4}{\pi^2} \left\{ \int_0^{\rho_{ab}} \text{arc sin } \frac{c_{cd}}{(c_{cc}\, c_{dd})^{\frac{1}{2}}} \frac{d\lambda_{ab}}{(1 - \lambda_{ab}^2)^{\frac{1}{2}}} \right.\right.$$

$$\left.\left. + \int_0^{\rho_{cd}} \text{arc sin } \frac{c_{ab}}{(c_{aa}\, c_{bb})^{\frac{1}{2}}} \frac{d\lambda_{cd}}{(1 - \lambda_{cd}^2)^{\frac{1}{2}}} + \begin{array}{l}\text{four similar integrals obtained} \\ \text{by interchanging } a, b, c, \text{ and } d\end{array} \right\} \right].$$

Plackett [5] has shown that, when the correlation coefficients are small, the sum of these six integrals is approximated by the expression

$$(\text{arc sin } \rho_{ab})(\text{arc sin } \rho_{cd})$$
$$+ (\text{arc sin } \rho_{ac})(\text{arc sin } \rho_{bd})$$
$$+ (\text{arc sin } \rho_{ad})(\text{arc sin } \rho_{bc}).$$

This may be compared with the corresponding term in (13). For the particular case in which $\rho_{13} = \rho_{14} = \rho_{24} = 0$, Plackett's approximation is closer than the approximation yielded by (13), even when ρ_{12}, ρ_{23}, ρ_{34} (any or all) are reasonably large. This is shown in Table II.

TABLE III

An Example Showing the Effect of Changing the Signs of the Correlation Coefficients in one of Schläfli's Examples [7] (*table originally given by Plackett* [5])

ρ_{12}	ρ_{23}	ρ_{34}	$2\sum\theta$	$4F$	P_4
1/2	1/2	1/2	1	2/15	2/15
1/2	−1/2	1/2	1/3	2/15	11/120
1/2	1/2	−1/2	1/3	−2/15	3/40
−1/2	1/2	−1/2	−1/3	2/15	1/20
+1/2	−1/2	−1/2	−1/3	−2/15	1/30
−1/2	−1/2	−1/2	−1	2/15	1/120

It is interesting to note than when $\rho_{12} = \rho_{23} = \rho_{34} = \rho_{14} = \frac{1}{2}$, $\rho_{13} = \rho_{24} = 0$ (line 19 on Table II), Plackett's approximation is exact. In light of the discussion to be given in Section 8, below, one might be led to conjecture that his approximation is optimum for the cases typified by this example.

8. An approximation for the particular case in which $\rho_{13} = \rho_{14} = \rho_{24} = 0$. Let P_4 be written as $P_4 = \frac{1}{16}[1 + 2\sum\theta + 4F]$ and consider what happens when the signs of any of the three non-vanishing correlation coefficients are changed. The example in Table III will suffice to illustrate.

Now, in (13) we have approximated F as $[\theta_{12}\theta_{34}/(1 + \theta_{23})]$, but the magnitude of F does not change with the sign of θ_{23}. Furthermore, if we let

$$F = [\theta_{12}\theta_{34}/(1 - |\theta_{23}|)], \tag{14}$$

the approximation becomes exact for this particular case, ($|\rho_{12}| = |\rho_{23}| = |\rho_{34}| = \frac{1}{2}$). The effect for the other examples is shown in the last column of Table II.

9. Conclusion. Equation (13) yields a satisfactory approximation for the normal multivariate integral. However, for special cases (such as for P_4 when $\rho_{13} = \rho_{14} = \rho_{24} = 0$, it is possible to obtain better approximations.

REFERENCES

1] Anis A. A. and Lloyd E. H. (1953). On the range of the partial sums of a finite number of independent normal variates. *Biometrika* **40** 35–42.

[2] Bacon Ralph H. (1958). Runs (or sequences) in correlated data. *J. Amer. Statist. Assoc.* **53** 569 (Abstract).

[3] Coxeter H. M. S. (1935). The functions of Schläfli and Lobatchefsky. *Quart. J. Math. Oxford Ser.* **6** 13–29.

[4] McFadden J. A. (1956). An approximation for the symmetric, quadrivariate normal integral. *Biometrika* **43** 206–207.

[5] Plackett R. L. (1954). A reduction formula for normal multivariate integrals. *Biometrika* **41** 351–360.

[6] Ruben H. (1954). On the moments of order statistics in samples from normal populations. *Biometrika* **41** 200–227.

[7] Schläfli L. (1860). On the multiple integral $\int^n dx\, dy \cdots dz$, whose limits are $\cdots$ (translated by Cayley, and published by him after Schläfli's death). *Quart. J. Pure Appl. Math.* **3** 97–108.

[8] Sondhi Man Mohan (1961). A note on the quadrivariate normal integral. *Biometrika* **48** 201–203.