

Integration of the General Bivariate Gaussian Distribution over an Offset Circle

By A. R. DiDonato and M. P. Jarnagin

1. Introduction. An efficient method is described in this paper for the numerical evaluation by a high-speed digital computer of the integral of any uncorrelated elliptical Gaussian distribution over the area of any arbitrarily centered circle in the plane. Actually, by introducing at most four linear transformations of coordinates, the present method can be used to cover the general problem of integrating any correlated or uncorrelated bivariate normal distribution over an arbitrary ellipse in the plane.

The basic integral to be evaluated is expressible in the form

$$(1) \quad P = \frac{1}{2\pi\sigma_x\sigma_y} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(r) \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] dx dy,$$

where the exponential represents an uncorrelated bivariate normal distribution centered at the origin with standard deviations σ_x and σ_y . A point target T is assumed at some arbitrary point (h, k) in the Oxy plane. The distance r is ordinary distance $[(x - h)^2 + (y - k)^2]^{1/2}$, where (x, y) is an arbitrary burst point in the plane. The function $p(r)$ by definition has the value one if r does not exceed a given constant R , and $p(r)$ is zero otherwise. This has the effect of limiting the field of integration to a circle with center at (h, k) and radius R . Hence the integral can be written as

$$(2) \quad P = \frac{1}{2\pi\sigma_x\sigma_y} \int_{h-R}^{h+R} \int_{k-\sqrt{R^2-(x-h)^2}}^{k+\sqrt{R^2-(x-h)^2}} \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] dx dy.$$

The problem was previously studied in [2], [3], and was mentioned in [4]. The more general problem of integrating an arbitrary integrable function of two variables over the circle has been analyzed in [5], [6] and other papers.

2. Transformations of the Basic Integral. The efficient numerical integration of the iterated integral of equation (2) is not a straightforward procedure for the following reasons:

(a) The behavior of the integrand in the second integration is such that for particular values of the input parameters it attains the shape of a spike with virtually all of its area contained within perhaps 1/100 or 1/1000 of the entire range of integration $[h - R, h + R]$;

(b) The slope with respect to x of the integrand in the second integration approaches infinity as x approaches $h + R$.

The minimum requirements for the ranges of values of the input variables were

$$\frac{1}{6} \leq \frac{\sigma_x}{\sigma_y} \leq 6, \quad 0 \leq \frac{h}{\sigma_x} \leq 20, \quad 0 \leq \frac{k}{\sigma_y} \leq 20.$$

Received October 27, 1960; revised February 10, 1961.

These were considered realistic ranges in the light of likely applications, for example, to problems of weapons assessment and other problems of operations research. The program which was developed, however, is valid for much wider ranges of values, as discussed in Section 8.

If a normalizing substitution

$$(3) \quad u = (x - h)/R$$

is made in the integral of equation (2), the range of integration on u is from -1 to 1 . If the resulting integral in u is split at zero and the sign of u is changed in the integral from -1 to 0 , the integration range can be reduced to the interval from zero to one. The difficulty of an unbounded slope of the integrand in the neighborhood of the upper limit of integration, here $u = 1$, is resolved by making another substitution,

$$(4) \quad t = \sqrt{1 - u}$$

yielding P in the following form:

$$(5) \quad P = \frac{R/\sigma_x}{\sqrt{2\pi}} \int_0^1 \left\{ \exp\left(-\frac{1}{2} \left[\frac{h - R(1 - t^2)}{\sigma_x} \right]^2\right) + \exp\left(-\frac{1}{2} \left[\frac{h + R(1 - t^2)}{\sigma_x} \right]^2\right) \right\} \{\text{Erf}(d_{11}) - \text{Erf}(d_{01})\} t dt$$

where

$$(6) \quad d_{11} = \frac{k + Rt\sqrt{2 - t^2}}{\sqrt{2}\sigma_y} \quad \text{Erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-v^2} dv$$

$$d_{01} = \frac{k - Rt\sqrt{2 - t^2}}{\sqrt{2}\sigma_y}$$

Equation (5) can be considered as a single integral if the $\text{Erf}(d_{11})$ and $\text{Erf}(d_{01})$ are numerically evaluated by analytic procedures rather than by quadratures. This is actually what is done. A table of the functions $\text{Erf}(x)$ and $\text{Erf}'(x)$ is assumed stored in the computer at suitably small intervals of the argument. The value of either of these functions, for any required value of the argument, is rapidly determined by a table look-up and the use of no more than four additional terms of the Taylor expansion about the nearest value for which the functions are stored. The truncation error, or remainder term, is then subject to a rigorous predetermined bound, as discussed of Section VI of [1]. The efficient methods of Hammer [5], Peirce [6], and others, for integration by quadrature over multi-dimensional domains, were considered at the request of the referee. It appears evident, however, that it is more efficient to take advantage of the opportunity to perform one of the two integrations analytically, since it can be done much more rapidly and systematically in this way than by any numerical quadrature method, and the other by Gaussian quadrature for single integrals. A small sample of test cases was run and this conclusion was in general verified. See Table 1.

3. Determination of Reduced Intervals of Integration. The difficulty described in the preceding section under (a) can be virtually eliminated by changing the

limits of integration on t to e_0 and e_1 , where $0 \leq e_0 \leq e_1 \leq 1$, and to a lesser degree by changing the limits of the error functions that occur in equation (5) from d_{01} and d_{11} to d_0 and d_1 , where $d_{01} \leq d_0 \leq d_1 \leq d_{11}$.

For a given ϵ in the range $0 < \epsilon < 1$ there is a unique rectangle, S , centered at the origin, with sides of lengths $2a(\epsilon)\sigma_x$ and $2a(\epsilon)\sigma_y$, which are parallel to the x - and y -axes, respectively (see Figure 1), such that

$$(7) \quad \bar{P} = \frac{1}{2\pi\sigma_x\sigma_y} \int_{-a\sigma_x}^{a\sigma_x} \int_{-a\sigma_y}^{a\sigma_y} \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] dx dy = 1 - \epsilon.$$

Transforming this integral by elementary methods, the following relation between ϵ and a can be derived

$$(8) \quad \text{Erf} \left(\frac{a}{\sqrt{2}} \right) = \sqrt{1 - \epsilon}.$$

All but ϵ of the entire bivariate distribution falls within the rectangle, by equation (7). Therefore, in carrying out the integration over the circle, as indicated by equation (5), only that part of the interior of the circle which also lies in the interior of the rectangle need be considered. If one integrates only over the common part of the interior of the circle with that of the rectangle (shaded area in Figure 1), denoting the result by P_1 , an error less than ϵ in magnitude will be committed. Thus the values of t can be restricted by the following inequalities

$$(9) \quad \frac{R - h - a\sigma_x}{R} \leq t^2 \leq \frac{R - h + a\sigma_x}{R}.$$

This can be seen by noting that the left and right boundaries of the rectangle are the lines $x = -a\sigma_x$ and $x = a\sigma_x$, and getting the corresponding value of t , using equations (3) and (4). But the entire integration interval on t in equation (5) is from zero to one. Hence the actual range of integration $[e_0, e_1]$ will be the common part of the interval, specified by (9), with the interval $[0, 1]$. The effect of this is

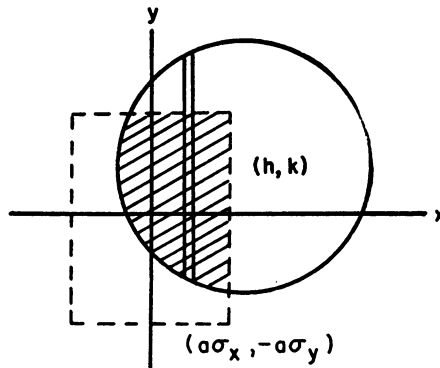


FIG. 1

that the limits e_0 and e_1 are given by

$$\begin{aligned}
 (10a) \quad & \left. \begin{aligned} e_1 = & \begin{cases} \sqrt{\frac{R - h + a\sigma_x}{R}} & \text{if } 0 < \frac{h - a\sigma_x}{R} < 1 \\ 1 & \text{if } \frac{h - a\sigma_x}{R} \leq 0 \\ 0 & \text{if } \frac{h - a\sigma_x}{R} \geq 1 \end{cases} \end{aligned} \right\} \\
 (10b) \quad & \\
 (10c) \quad & \\
 (10d) \quad & \left. \begin{aligned} e_0 = & \begin{cases} \sqrt{\frac{R - h - a\sigma_x}{R}} & \text{if } \frac{h + a\sigma_x}{R} < 1 \\ 0 & \text{if } \frac{h + a\sigma_x}{R} \geq 1 \end{cases} \end{aligned} \right\} \\
 (10e) \quad &
 \end{aligned}$$

In order to keep the interval over which the integration is carried out as small as possible, if the analogous difference $e_1' - e_0'$ for σ_y and k is smaller than the difference $e_1 - e_0$ based on σ_x and h, x, σ_x , and h are interchanged with y, σ_y , and k respectively. This interchange is always possible because of the similar roles played in equation (1) by x and y and their associated parameters.

The integration interval $[d_{01}, d_{11}]$ of equation (5) can also be reduced in many cases by analysis similar to that employed above for $[e_0, e_1]$. The new limits will be defined as follows:

$$\begin{aligned}
 (11a) \quad & \left. \begin{aligned} d_1 = & \begin{cases} d_{11} & \text{if } d_{11} < \frac{a(\epsilon)}{\sqrt{2}} \\ \frac{a(\epsilon^*)}{\sqrt{2}} & \text{if } d_{11} \geq \frac{a(\epsilon)}{\sqrt{2}} \end{cases} \end{aligned} \right\} \\
 (11b) \quad & \\
 (12a) \quad & \left. \begin{aligned} d_0 = & \begin{cases} d_{01} & \text{if } d_{01} > -\frac{a(\epsilon)}{\sqrt{2}} \\ -\frac{a(\epsilon^*)}{\sqrt{2}} & \text{if } d_{01} \leq -\frac{a(\epsilon)}{\sqrt{2}} \end{cases} \end{aligned} \right\} \\
 (12b) \quad &
 \end{aligned}$$

where

$$\epsilon^* = \frac{1}{2}\epsilon - \frac{1}{18}\epsilon^2.$$

Further, since, by equation (8), $\text{Erf} [a(\epsilon^*)/\sqrt{2}] = \sqrt{1 - \epsilon^*} = 1 - \frac{1}{4}\epsilon$, $\text{Erf} (d_{11})$ in equation (5) can be replaced by $(1 - \epsilon/4)$ if (11b) holds, and similarly $\text{Erf} (d_{01})$ can be replaced by $(\epsilon/4 - 1)$ if (12b) holds. Moreover, because of the monotonic character of d_{11} and d_{01} as functions of t (see equations (6)), increasing and decreasing respectively on the interval $[0, 1]$, if (11b) holds for a particular value of t , it will hold for all subsequent values of t in the integration, and similarly for d_{01} and (12b). A detailed analysis is given in Appendix A of [1].

Thus the probability P as given by equation (5) is approximated within an

arbitrarily chosen ϵ by P_1 , where

$$(13a) \quad P_1 = \frac{R/\sigma_x}{2\sqrt{2}} \int_{e_0}^{\epsilon_1} (f_1 + f_2)[\text{Erf}(d_1) - \text{Erf}(d_0)] t dt$$

$$(13b) \quad f \equiv f_1 + f_2 \equiv \frac{2}{\sqrt{\pi}} \exp\left(-\frac{1}{2} \left[\frac{h - R(1 - t^2)}{\sigma_x}\right]^2\right) + \frac{2}{\sqrt{\pi}} \exp\left(-\frac{1}{2} \left[\frac{h + R(1 - t^2)}{\sigma_x}\right]^2\right).$$

4. Determination of Conditions for $P < 5\epsilon$, $P > 1 - 5\epsilon$. In order to improve the efficiency of the machine computing program, conditions were determined such that, if the kill probability were less than 5ϵ , a result of zero would be stored without further computation, or a result of one would be stored if P were greater than $1 - 5\epsilon$. The details are given in [1].

5. Determination of the Gaussian Order of Integration. The method chosen for evaluating equation (13a) numerically is that of Gaussian quadrature. The symbol $O(G)$ will denote the order of Gaussian multipliers. The orders stored in the NORC program are 6, 8, 12, 16, 20, and 24 [7]. For the ranges of values of the input parameters for which the program is designed, it is never necessary to use more than 16(24) integration points over the entire interval of integration in order to obtain results correct to three (six) decimal digits.

The rule for determining $O(G)$ in the program is empirical, but it was developed by an extensive program of test cases, with adequate checking of results, as discussed in Section 9.

Two levels of accuracy were considered, three and six decimal digits for the computed probabilities. For each accuracy level, $O(G)$ depends on the magnitude of a positive quantity N given by the empirical equation

$$(14) \quad N = (e_1 - e_0) \left[0.17 R/\sigma_x + \frac{R/\sigma_y}{.05 |R - k|/\sigma_y + 10} \right].$$

In the program for six-digit accuracy ($\epsilon = 10^{-7}$), $O(G)$ is given by the following rules:

- If $N \geq 2.75$, $O(G) = 24$; if $1.4 \leq N < 2.75$, $O(G) = 20$;
- if $0.8 \leq N < 1.4$, $O(G) = 16$; if $0.4 \leq N < 0.8$, $O(G) = 12$;
- if $0.15 \leq N < 0.4$, $O(G) = 8$; if $N < 0.15$, $O(G) = 6$.

In the program for three-digit accuracy ($\epsilon = 10^{-4}$):

- If $N \geq 2.0$, $O(G) = 16$; if $0.8 \leq N < 2.0$, $O(G) = 12$;
- if $0.5 \leq N < 0.8$, $O(G) = 8$; if $N < 0.5$, $O(G) = 6$.

6. Computation of the Error Function $\text{Erf}(x)$ and Its Derivative. The required values of $\text{Erf}(x)$ and its derivative, $(2/\sqrt{\pi})e^{-x^2}$, are obtained with a subroutine

based on Taylor series expansions [8] about stored values of these functions, and taking account of theoretical bounds on the error terms in the series. The numerical values of the error function and its derivative are stored at intervals of 0.01 from $x = 0$ to $x = 3.85$. Therefore this subroutine is very fast in that it requires only two terms of the Taylor series for four-digit accuracy in the computed functions, and only four terms for nine-digit accuracy, including in each case the rapidly accessible stored value. These accuracy levels are acceptable in the programs for three and six decimal digit computed probabilities, respectively, as discussed in detail in Section VII of [1].

7. Error Analysis. A detailed error analysis of the computations is omitted here. However, it is shown in Section VII of [1] that for every case within the permissible ranges of the input parameters, the total error in the computed probability, due to round-off, truncation of series, and all other causes, is less than 5ϵ . The quantity ϵ is taken as 10^{-4} and 10^{-7} in the programs for three and six decimal digit accuracy respectively.

8. Nature and Scope of the Computing Program. The program is constructed so as to accept the five input parameters R , σ_x , σ_y , h , k , even though there exist only four independent combinations. The limitations in the permissible ranges for the present program are

$$(15) \quad \begin{aligned} \frac{1}{15} &\leq \sigma_x/\sigma_y \leq 15; \\ 0 &\leq h/\sigma_x \leq 600; \\ 0 &\leq k/\sigma_y \leq 600; \\ R &> 0. \end{aligned}$$

These ranges greatly exceed the originally contemplated requirements (Section 2 above). The primary reasons for this extension were to analyze the function P thoroughly and to allow for possible extensions of the realistic ranges in the future.

9. Discussion of Results. Since the order $O(G)$ of Gaussian multipliers for a given case is determined by an empirical rule (Section 5), it was necessary to check the validity of the rule with an extensive test program of cases spanning the space of permissible values of the input parameters. More than 7,000 cases were computed in this test program, and the accuracy of the results was verified by doubling the number of integration points, and also in borderline cases by checking against values computed by an independent iterative Simpson's rule routine. In addition, the tables of values given in [2] and [3] were computed independently by the present program. The results agreed, within one unit in the last place retained by the respective authors, except for one discrepancy in the case of each author, where the difference was more than one unit. In these two cases, as well as in all other cases of the total of more than 7,000, it was established that the results given by the present program were correct to the specified extent, three or six decimal digits.

The input values used in the test program are given in Table III of [1]. Briefly, $\sigma_x = 1$, $\sigma_y = 1, 3, 6, 10, 15$; h and k range from 0 to 600. For a given set of values

TABLE 1

Case No.	R	σ_x	σ_y	h	k	O(G)	(D & J) $P \times 10^3$	(H & P) $P \times 10^3$	(%) Error
1	6.6282	1	3	2	0.2	16	960.2043	958.2172	-.207
2	8.3092	1	3	2	0.2	16	991.9049	1045.322	+5.38
3	6.5918	1	4	2	0.2	16	865.1445	872.8980	+.896
4	8.0172	6	1	10	5	16	257.3936	258.3941	+.389
5	7.1589	10	1	10	5	16	236.2148	234.0224	-.928
6	0.14929	1	1	0	0	6	11.08189	11.08189	0
7	2.5205	1	1	0	0	12	958.2668	958.2629	-.0004
8	2.9999	1	1	0	0	12	988.8877	988.8369	-.0051
9	2.7705	1	1	0	2	12	709.6000	709.5979	-.0003
10	3.4633	1	1	0	2	16	898.1620	898.1085	-.0059
11	4.2076	1	1	0	2	16	979.1231	978.5897	-.0544
12	4.9999	1	1	0	2	16	997.7784	996.0434	-.1739
13	2.3400	1	1	4	4	12	2784172	2783833	-.0122
14	2.9588	1	1	4	4	12	2.405112	2.403328	-.0742
15	4.5226	1	1	4	4	12	108.6807	107.7243	-.8800
16	6.4408	1	1	4	4	20	758.2833	732.6466	-3.3809
17	8.6568	1	1	4	4	20	998.2986	981.1265	-1.7201
18	1.5754	1	3	2.5	2	8	38.25434	38.25434	0
19	3.8489	1	3	2.5	2	12	501.7925	501.8157	+.0046
20	6.7930	1	3	2.5	2	16	914.5639	898.7277	-1.7315
21	8.6617	1	6	4	20	16	19.75531	19.82685	+.3621
22	1.7024	1	3	1.5	5	8	52.70654	52.70654	0
23	4.3806	1	3	1.5	5	16	365.2462	365.2257	-.0056
24	3.1229	1	3	1.5	10	12	7.033516	7.033444	-.0010
25	4.6891	1	3	1.5	10	16	29.99226	30.03412	+.1406
26	6.5438	1	3	1.5	10	16	108.7028	112.7367	+3.7109
27	4.1553	1	6	0.75	3	12	439.3769	439.5179	+.0321
28	7.8804	1	6	4	3	16	675.0670	692.2308	+2.5425
29	4.000	1	1/2	0.86	1.3	20	997.6353	967.9286	-2.977
30	4.000	1	1/2	2.58	1.3	20	873.9654	840.5547	-3.822
31	4.000	1	1/2	3.44	1.3	20	618.3736	573.4521	-7.264
32	4.000	1	1/2	4.3	1.3	20	292.9237	267.4685	-8.690
33	4.000	1	1/2	5.16	1.3	20	81.88281	76.40502	-6.689
34	8.000	8	1	17.20	4.29	16	92.76734	96.52873	+4.054
35	8.000	8	1	20.64	4.29	16	40.28306	42.02923	+4.334
36	8.000	8	1	24.08	4.29	16	14.85987	15.54811	+4.632
37	8.000	4	1	0	4.29	16	897.8378	925.2863	+3.057
38	8.000	4	1	4.64	4.29	16	686.5311	708.9009	+3.258
39	8.000	4	1	6.96	4.29	16	468.7122	486.4106	+3.775
40	8.000	4	1	9.28	4.29	16	257.9851	269.8771	+4.609
41	8.000	4	1	11.60	4.29	16	110.8666	117.0092	+5.540
42	8.000	4	1	13.92	4.29	16	36.33885	38.62267	+.6284
43	8.000	4	1	16.24	4.29	16	8.937987	9.528004	+6.601
44	10.279	1	3	2.5	2	16	995.6620	1096.646	+10.142
45	12.201	1	3	2.5	2	16	999.4961	1236.929	+23.755

D & J → Computation of P with DiDonato-Jarnagin procedure.

H & P → Computation of P with Hammer-Peirce procedure.

O(G) → Order of Gaussian multipliers used for D & J procedure.

All H & P results were computed with H & P multipliers of order 64.

$$\% \text{ error} = \frac{H \ \& \ P \ (\text{result}) - D \ \& \ J \ (\text{result})}{D \ \& \ J \ (\text{result})}$$

of σ_y , h and k , ten values of R were determined and used in such a way that the corresponding values of P ranged from roughly 10^{-4} to $1 - 10^{-5}$.

A table of 10,080 entries is given in Appendix C of [1]. This is an inverse table giving R as a function of P , $\sigma_x (=1)$, σ_y , h , k . Six values of P are used, .05, .20, .50, .70, .90, .95. The values of R were computed such that the specified probabilities, P , were obtained to within three or four units in the fifth significant digit.

U. S. Naval Weapons Laboratory
Dahlgren, Virginia

1. A. R. DIDONATO & M. P. JARNAGIN, *Integration of the General Bivariate Gaussian Distribution over an Offset Ellipse*, NWL Report 1710, U. S. Naval Weapons Laboratory, Dahlgren, Virginia, 11 August 1960, unclassified.
2. H. H. GERMOND, *Integration of the Gaussian Distribution over an Offset Ellipse*, RAND Corporation Report No. P-94, 28 July 1949, unclassified.
3. J. R. LOWE, *Integration of the Binormal Distribution over an Offset Circle*, ARDE Memorandum (B) 5/60, Armament Research and Development Establishment, Fort Halstead, Kent, England, February 1960, unclassified.
4. OEG Study 626, *Probability-of-Damage Problems of Frequent Occurrence*, Operations Evaluation Group, Office of the Chief of Naval Operations, 11 December 1959, unclassified.
5. P. C. HAMMER, "Numerical evaluation of multiple integrals," *On Numerical Approximation*, University of Wisconsin Press, Madison, Wisconsin, 1959, p. 99-115.
6. W. H. PEIRCE, "Numerical integration over the planar annulus," *J. Soc. Indust. Appl. Math.*, v. 5, 1957, p. 66-73.
7. (a) P. DAVIS & P. RABINOWITZ, "Abcissae and weights for Gaussian quadratures of high order," *J. Res. Nat. Bur. Standards*, v. 56, January 1956, p. 35 (for orders 20, 24).
(b) A. N. LOWAN, N. DAVIDS, & A. LEVENSON, "Table of the zeros of the Legendre polynomials of order 1-16 and the weight coefficients for Gauss' mechanical quadrature formula," p. 185-189 of *Tables of Functions and of Zeros of Functions*, Nat. Bur. Standards Appl. Math. Ser., No. 37, U. S. Government Printing Office, Washington, D.C., 1954 (for orders 6, 8, 12, 16)
8. NAT. BUR. STANDARDS APPL. MATH. SER. No. 41, *Tables of the Error Function and its Derivative*, U. S. Government Printing Office, Washington, D. C., 1954.