REMARKS ON COMPUTING THE PROBABILITY INTEGRAL IN ONE AND TWO DIMENSIONS

G. PÓLYA

STANFORD UNIVERSITY

Introduction

In the first part of the present paper the probability integral in one dimension is considered. This first part may be regarded as an illustration of the principle that "no problem whatever is solved completely." In fact, the problem of computing the total area under the Gaussian curve was solved by Laplace or, even before him, under a slightly different form, by Euler, and this solution has been presented since under various forms. In the present paper there are offered two different solutions of the same problem, and an inequality, derived from Laplace's solution, all of which seem to be new and are certainly very little known.

In the second part of the paper the probability integral in two dimensions is considered. In this part, which had its origin in a practical problem, formulas and inequalities which appear to be useful in computing volumes under the normal probability surface are presented.

We use the following notation:

$$g(x) = (2\pi)^{-1/2} e^{-x^2/2}, \tag{1}$$

$$G(x) = \int_0^x g(t)dt, \tag{2}$$

$$\begin{aligned} L &= L(a,a';b,b';r) \\ &= \int_a^b \int_{a'}^{b'} [2\pi(1-r^2)]^{-1/2} g\{[(x^2 - 2rxx' + x'^2)/(1-r^2)]^{1/2}\} dx'dx, \end{aligned} \tag{3}$$

$$M = M(h,k;r) = L(h,k;+\infty,+\infty;r). \tag{4}$$

The symbols $g(x)$ and $G(x)$ should remind us of "Gauss." The limits a,b in L correspond to x, and a',b' to x'. The quantity M is represented by an integral of the type

$$\int_h^\infty \int_k^\infty.$$

I. The Probability Integral in One Dimension

1. An inequality

We try to see something new in the most usual method of evaluating the total area under the Gaussian curve. Following that method, we consider

$$2G(a) = \int_{-a}^{a} (2\pi)^{-1/2} e^{-x^2/2} dx = \int_{-a}^{a} (2\pi)^{-1/2} e^{-y^2/2} dy ,$$

where a is any positive quantity. Therefore

$$(1.1) \qquad 2\pi[2G(a)]^2 = \int_{-a}^{a}\int_{-a}^{a} e^{-(x^2+y^2)/2} dxdy .$$

The integral (1.1) is extended over a square with area $4a^2$. Generalizing (1.1), we consider the integral

$$(1.2) \qquad \iint_R e^{-(x^2+y^2)/2} \, dxdy ,$$

where R is any region the area of which is $4a^2$. Specializing R, we consider a circular region, with center at the origin, of which the area is $4a^2$. We call this region R_0; its boundary has the equation

$$(1.3) \qquad \pi(x^2 + y^2) = 4a^2 .$$

The circle (1.3) is a closed level line of the integrand of (1.2). Let us vary R, subject only to the condition that its area remains constant, equal to $4a^2$. Then the integral (1.2) varies; we say that it attains its maximum when $R = R_0$.

In fact, the contribution of the common part of the two regions R_0 and R to

$$(1.4) \qquad \iint_{R_0} e^{-(x^2+y^2)/2} dx\, dy - \iint_R e^{-(x^2+y^2)/2} dx\, dy$$

is zero. Therefore it is sufficient to extend the first integral under (1.4) over that portion of R_0 which is not contained in R, and the second integral over that portion of R which is not contained in R_0. The areas of the two considered partial regions are equal, but in the former, which is inside the circle (1.3), the values of the integrand are greater than in the latter, which is outside (1.3). Hence the difference (1.4) is positive (unless R_0 coincides with R), which is the fact we wished to prove.[1]

The domain of integration in (1.1) is a special region R. Hence we infer that

$$2\pi[2G(a)]^2 < \iint_{R_0} e^{-(x^2+y^2)/2} dx\, dy = \int_0^{2\pi}\int_0^{2a\pi^{-1/2}} e^{-r^2/2} r\, dr\, d\varphi$$
$$= 2\pi(1 - e^{-2a^2/\pi}).$$

We note the result, writing x for a; for $x > 0$

$$(1.5) \qquad 2G(x) < (1 - e^{-2x^2/\pi})^{1/2}.$$

[1] This kind of argument is well known because of the importance given it in the work of J. Neyman and E. S. Pearson.

We can of course extend the foregoing argument from 2 to n dimensions. We note just the result: If we define

$$(1.6) \qquad a_n = 2\pi^{-1/2}[\Gamma\{(n/2)+1\}]^{1/n},$$

then we have

$$(1.7) \qquad G(x) < a_n^{-1}(2\pi)^{-1/2}\left[n\int_0^{a_n x} t^{n-1}e^{-t^2/2}\,dt\right]^{1/n}$$

for $n = 2, 3, \cdots$ and $x > 0$. Our (1.5) is the special case $n = 2$.

2. Other proofs of the inequality

We shall use (in the present section only) the symbol $H(x)$ defined by the equation

$$(2.1) \qquad 2H(x) = (1 - e^{-2x^2/\pi})^{1/2}.$$

With this abbreviation, we can write the inequality (1.5) in the concise form

$$(2.2) \qquad G(x) < H(x).$$

Comparing the sides of this inequality in various manners, we find that *all three functions*

$$(2.3) \qquad H(x) - G(x), \qquad H^2(x) - G^2(x), \qquad H(x)/G(x)$$

behave in the same way in the interval $0 < x < \infty$. *They first increase, reach a unique maximum, then decrease, and, as* $x \to \infty$, *each tends toward the same value that it takes at the point* $x = 0$.

The value taken at the two extremities of the interval $(0, \infty)$ is zero for the first two functions (2.3) and one for the last function. Therefore the proposition just stated involves that

$$(2.4) \quad H(x) - G(x) > 0, \qquad H^2(x) - G^2(x) > 0, \qquad H(x)/G(x) > 1$$

for $x > 0$. Thus we obtain three different proofs of the inequality (2.2), that is, of (1.5).

As x approaches infinity, both $G(x)$ and $H(x)$ tend to the same value 1/2. Moreover, we have the expansions into powers of x

$$G(x) = (2\pi)^{-1/2}x\left[1 - \frac{x^2}{6} + \cdots\right],$$

$$H(x) = (2\pi)^{-1/2}x\left[1 - \frac{x^2}{2\pi} + \cdots\right]$$

which show that $H(x)$ is greater than $G(x)$ for small positive values of x. Both functions take the value zero at the point $x = 0$, and their quotient takes the value one there.

These remarks prove a certain part of the proposition stated. We have still to show, however, that the derivatives of the three functions (2.3) behave in the same way in the interval $x > 0$: they change sign only once, *each derivative has just one positive root which is simple.*

Since it would take up too much space to discuss all three cases, we shall restrict ourselves to the last function, which is the most instructive. The equation

$$[H(x)/G(x)]' = 0$$

can be written in the form

$$\frac{\pi}{2x}(e^{2x^2/\pi} - 1) - e^{x^2/2}\int_0^x e^{-t^2/2}\,dt = 0. \tag{2.5}$$

We need the expansion of the left-hand side into powers of x. We consider, therefore, the function

$$y = e^{x^2/2}\int_0^x e^{-t^2/2}\,dt \tag{2.6}$$

and observe that it vanishes at the point $x = 0$ and satisfies the differential equation

$$y' = xy + 1. \tag{2.7}$$

We expand the solution of this equation into powers of x by the usual method, taking into account the initial condition that $y = 0$ as $x = 0$. Thus we obtain the elegant series

$$e^{x^2/2}\int_0^x e^{-t^2/2}\,dt = \frac{x}{1} + \frac{x^3}{1.3} + \frac{x^5}{1.3.5} + \frac{x^7}{1.3.5.7} + \cdots. \tag{2.8}$$

Now we can write (2.5), expanding the left-hand side, in the form

$$\sum_{n=2}^{\infty}\left[\frac{1}{1.2.3\cdots n}\left(\frac{2}{\pi}\right)^{n-1} - \frac{1}{1.3.5\cdots(2n-1)}\right]x^{2n-1} = 0. \tag{2.9}$$

In this expansion we say that the coefficient of x^3 is negative and the other coefficients, those of $x^5, x^7, \cdots$, are positive. For the first two coefficients (those of x^3 and x^5) this assertion may be verified by computation. For the other coefficients we use mathematical induction. If we know that

$$\frac{1.3.5\cdots(2n-1)}{1.2.3\cdots n}\left(\frac{2}{\pi}\right)^{n-1} > 1$$

and $n \geqq 3$, we can infer that

$$\frac{1.3\cdots(2n-1)(2n+1)}{1.2\cdots n(n+1)}\left(\frac{2}{\pi}\right)^{n} > \frac{2n+1}{n+1}\frac{2}{\pi} \geqq \frac{7}{4}\frac{2}{\pi} > 1.$$

Dividing the left-hand side of (2.9) by x^5, we obtain a series of the form

$$-\frac{a}{x^2} + a_0 + a_1x^2 + a_2x^4 + \cdots, \tag{2.10}$$

where all the numbers $a, a_0, a_1, a_2, \cdots$ are positive. Now, obviously, a series of the form (2.8) represents a steadily increasing function, varying from $-\infty$ to ∞ as x increases from zero to infinity. Such a function takes the value zero just once, and so we have proved our assertion.[2]

3. An approximation to the probability integral

On the basis of the proposition proved in the foregoing section, the usual tables allow a quick evaluation of the unique maximum of $H(x)/G(x)$. Thus we find that the following remark may be made concerning (1.5): If we take the right-hand side of this inequality as an approximation to $2G(x)$, the error committed is less than one per cent (even less than 0.71 per cent) of the quantity approximated.

4. A derivation of the total area

We may rewrite (2.8) in the form

$$\int_0^x e^{-t^2/2}\,dt = \frac{\dfrac{x}{1} + \dfrac{x^3}{1.3} + \cdots + \dfrac{x^{2n-1}}{1.3.5 \cdots (2n-1)} + \cdots}{1 + \dfrac{x^2}{2} + \dfrac{x^4}{2.4} + \cdots + \dfrac{x^{2n}}{2.4.6 \cdots 2n} + \cdots}. \tag{4.1}$$

For any positive value of x, there is, in each series on the right-hand side, a term whose absolute value is maximum. Being given x, we locate the maximum term by examining the quotient of the general term and of the foregoing term, which is

$$\frac{x^2}{2n-1} \quad \text{and} \quad \frac{x^2}{2n}$$

in the numerator and in the denominator, respectively. We examine especially the place where this quotient passes the value one, and we find the following: When

$$\sqrt{2n} < x < \sqrt{2n+1}, \tag{4.2}$$

the maximum term in the numerator is the one containing x^{2n-1}, and the maximum term in the denominator is the one containing x^{2n}. Now we can foresee heuristically, and confirm afterward by rigorous argument, that *the quotient of the series differs from the quotient of the maximum terms only by a quantity*

[2] We have proved here, in fact, a simple case of an extension of Descartes' rule of signs. See G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis* (1925), vol. 2, p. 43, problems 38 and 40.

tending to zero when x *tends to infinity*. Thus, connecting x and n by (4.2), we obtain from (4.1) that

$$\lim_{x\to\infty} \int_0^x e^{-t^2/2}\, dt = \lim_{x\to\infty} \frac{\dfrac{x^{2n-1}}{1.3.5 \cdots (2n-1)}}{\dfrac{x^{2n}}{2.4.6 \cdots 2n}} \tag{4.3}$$

$$= \lim_{n\to\infty} \frac{2.4.6 \cdots 2n}{1.3.5 \cdots (2n-1)} \frac{1}{(2n)^{1/2}} = \left(\frac{\pi}{2}\right)^{1/2}.$$

The last equation is based on Wallis' formula. We have found a new derivation of the total area under the probability curve.

The heuristic reasoning suggested above is as follows: In each series, the terms "close" to the maximum term differ but little from it, and the terms which are "far" from it are relatively small so that they contribute only a negligible amount to the sum of the series. Thus the proportion of the maximum terms tends to become the proportion of the series. This heuristic reasoning contains a certain germ from which the rigorous proof can be evolved.[3]

The value of the probability integral can be derived from Wallis' formula in various ways, all rather different from the foregoing.[4]

5. Another derivation of the total area

Many definite integrals can be evaluated by means of the calculus of residues, but the one expressing the total area under the normal curve is not so evaluated in the usual textbooks.[5] We shall see, however, that such an evaluation is possible.[6]

We work now in the complex z-plane. We call P the parallelogram with vertices

$$R + iR, \qquad -R - iR, \qquad -R + 1 - iR, \qquad R + 1 + iR;$$

R is positive (and large). The center of P is at the point $z = 1/2$, two sides of P are horizontal and of length one, the other two sides pass through the points $z = 0$ and $z = 1$, and the side passing through the origin bisects the angle between the coördiante axes. We take the integral

$$\oint e^{\pi i z^2} \tan \pi z \cdot dz \tag{5.1}$$

counterclockwise around the boundary of P. The integrand has just one

[3] The rigorous proof follows, as a special case, from G. Pólya and G. Szegö, *op. cit.*, vol. 2, p. 12, problem 72 (with $\beta = 1$, $b = 1/2$, $k = 1/2$).

[4] See especially T. J. Stieltjes, *Œuvres complètes*, vol. 2, pp. 263–264.

[5] See G. N. Watson, *Complex Integration and Cauchy's Theorem*, Cambridge Tracts No. 15 (1914), p. 79.

[6] This will not surprise anyone who is familiar with the evaluation of the Gaussian sums, important in the theory of numbers, by means of complex integration. The argument used there yields, as will be shown here, the desired definite integral, if used in the opposite direction in an appropriate special case.

singular point inside P, a simple pole at the point $z = 1/2$, the center of P, and the residue at this pole is

$$\frac{e^{\pi i/4} \cdot 1}{-\pi}. \tag{5.2}$$

Thus we know the value of (5.1). We let R tend to infinity. Then the contribution of the two horizontal sides becomes negligible, and (5.1) goes over into two integrals, extended along infinite parallel straight lines. We can transform the integral along the right-hand line by a change of the variable of integration, and so we finally obtain

$$\int_{-(1+i)\infty}^{(1+i)\infty} [e^{\pi i(z+1)^2} - e^{\pi iz^2}] \tan \pi z \cdot dz = 2\pi i e^{\pi i/4}/(-\pi)\,.$$

The integral is extended along the first bisector of the coördinate axes. Hence we obtain, transforming the integrand,

$$\begin{aligned} 2e^{\pi i/4}/i &= \int_{-(1+i)\infty}^{(1+i)\infty} (-e^{\pi iz^2+2\pi iz} - e^{\pi iz^2}) \frac{e^{2\pi iz} - 1}{e^{2\pi iz} + 1} \frac{dz}{i} \\ &= (1/i)\int_{-(1+i)\infty}^{(1+i)\infty} e^{\pi iz^2} (1 - e^{2\pi iz}) dz \\ &= (1/i)\left[\int_{-(1+i)\infty}^{(1+i)\infty} e^{\pi iz^2} dz + \int_{-(1+i)\infty}^{(1+i)\infty} e^{\pi i(z+1)^2} dz\right] \\ &= (1/i)\left[\int_{-(1+i)\infty}^{(1+i)\infty} e^{\pi iz^2} dz + \int_{1-(1+i)\infty}^{1+(1+i)\infty} e^{\pi iz^2} dz\right] \\ &= (2/i)\int_{-(1+i)\infty}^{(1+i)\infty} e^{\pi iz^2} dz\,. \end{aligned} \tag{5.3}$$

The step before the last one was a change of the variable of integration; it involves shifting the line of integration by a unit length westward. The last step shifts back this line of integration until it coincides with the first bisector of the axes. This step is justified by Cauchy's theorem and involves, if considered in detail, integration around the parallelogram P and shifting the horizontal sides of P to infinity, as before.

Finally we change the variable of integration in the last integral under (5.3), setting

$$z = e^{\pi i/4}\, t\,.$$

Then t is real, and we obtain

$$2e^{\pi i/4}/i = (2e^{\pi i/4}/i)\int_{-\infty}^{\infty} e^{-\pi t^2}\, dt\,.$$

This equation yields, after a trivial transformation, the total area under the normal curve.

II. The Probability Integral in Two Dimensions

6. Results

We consider now the volume under the normal bivariate surface and over a rectangle whose sides are parallel to the axes. The numerical value of this volume, denoted by $L(a,a';\, b,b';\, r)$ in the introduction under (3), is often needed in practical applications of statistics. The computation of L is easily reduced to that of $M(h,k;\, r)$, defined under (4), and there exist well-known tables for M (*Tables for Statisticians and Biometricians*, Part II). These tables, however, demand a rather unsatisfactory interpolation, and therefore it may be advantageous to use numerical integration or some other kind of approximate formulas instead of, or conjointly with, the tables. Various results helpful in such work have been obtained recently in connection with a practical problem.[7] Three of them will be stated here: a double inequality, an approximate formula, and an enveloping series. The proofs of these results will be given in the last three sections.

a) A double inequality.—We suppose that

$$0<r<1, \qquad rh-k>0. \tag{6.1}$$

Under these conditions

$$M(h,k;r)<\frac{1}{2}-G(h), \tag{6.2}$$

$$M(h,k;r)>\frac{1}{2}-G(h)-\frac{1-r^2}{rh-k}g(k)\left[\frac{1}{2}-G\left(\frac{h-rk}{(1-r^2)^{1/2}}\right)\right]. \tag{6.3}$$

The abbreviations g, G, and M are defined in the introduction under (1), (2), and (4). The right-hand side of the rather obvious inequality (6.2) is the limit toward which the left-hand side tends when h and r are constant and k approaches $-\infty$; or when h and k are constant and r approaches one. This follows from (6.3).

b) An approximate formula.—We assume that $a<b$, $a'<b'$, and that r differs but little from one. We define nine new quantities α, α', β, β', γ, γ', δ, δ', and ρ by the equations

$$\rho=\left(\frac{1-r}{1+r}\right)^{1/2}, \tag{6.4}$$

$$\alpha,\ \beta,\ \gamma,\ \delta=\frac{a+a'}{[2(1+r)]^{1/2}},\ \frac{b+b'}{[2(1+r)]^{1/2}},\ \frac{b+a'}{[2(1+r)]^{1/2}},\ \frac{a+b'}{[2(1+r)]^{1/2}}, \tag{6.5}$$

$$\alpha',\ \beta',\ \gamma',\ \delta'=\frac{-a+a'}{[2(1-r)]^{1/2}},\ \frac{-b+b'}{[2(1-r)]^{1/2}},\ \frac{-b+a'}{[2(1-r)]^{1/2}},\ \frac{-a+b'}{[2(1-r)]^{1/2}}. \tag{6.6}$$

[7] In connection with the same project, tables have been computed in the Statistical Laboratory, University of California, by Leo A. Aroian, E. Fix, and Madeline Johnsen. It is hoped that they will be published in the near future.

It is understood that in (6.5) and (6.6) the first quantity on the left-hand side is equal to the first one on the right-hand side, the second on the left to the second on the right, and so on. With these abbreviations, we have the approximate formula:

$$(6.7)\quad \begin{aligned} L(a,a';b,b';r) \sim & [G(\beta) - G(\alpha)][G(\delta') - G(\gamma')] \\ & + \rho g(\alpha)[g(\gamma') - 2g(\alpha') + g(\delta') + \alpha'\{G(\gamma') - 2G(\alpha') + G(\delta')\}] \\ & + \rho g(\beta)[g(\gamma') - 2g(\beta') + g(\delta') + \beta'\{G(\gamma') - 2G(\beta') + G(\delta')\}] . \end{aligned}$$

This approximation must be used with caution, but it turned out to be especially useful in the particular case in which

$$a = a' , \qquad b = b' .$$

These relations involve

$$\gamma = \delta = \frac{\alpha + \beta}{2} , \qquad \alpha' = \beta' = 0 , \qquad \gamma' = -\delta'$$

and (6.7) takes the much simpler form

$$(6.8)\quad \begin{aligned} L(a,a;b,b;r) \sim & 2[G(\beta) - G(\alpha)]G(\delta') \\ & - 2\rho[g(\alpha) + g(\beta)][g(0) - g(\delta')] . \end{aligned}$$

c) *An enveloping series.*—If an infinite series

$$(6.9)\quad a_0 - a_1 + a_2 - a_3 + \cdots$$

and a number A are so related that

$$(6.10)\quad A < a_0 , \qquad A > a_0 - a_1 , \qquad A < a_0 - a_1 + a_2 , \cdots ,$$

then we say that the series (6.9) *envelopes* A, and we write

$$(6.11)\quad A \propto a_0 - a_1 + a_2 - a_3 + \cdots .$$

Thus the specific use of the symbol $\propto$ in (6.11) expresses an infinite system of inequalities, the inequalities (6.10), which we could also express by saying that A is contained between any two consecutive partial sums of the series (6.9). An enveloping series may be divergent; if it is convergent the enveloped number is its sum.[8]

There is a well-known divergent enveloping series connected with the probability integral in one dimension:

$$(6.12)\quad \int_x^\infty g(t)\,dt \propto g(x)\left[\frac{1}{x} - \frac{1}{x^3} + \frac{1 \cdot 3}{x^5} - \frac{1 \cdot 3 \cdot 5}{x^7} + \cdots\right] ;$$

x is assumed to be positive. It may be deserving of some interest that there is a series of similar nature connected with the probability integral in two dimen-

[8] For terminology and examples, see G. Pólya and G. Szegö, *op. cit.*, vol. 1, pp. 26–29.

sions. We consider, following C. Nicholson,[9] $V(h,k)$, the volume under the normal bivariate surface with $r = 0$ and over a right triangle with vertices (0,0), $(h,0)$, (h,k):

$$V(h,k) = \int_0^h \int_0^{kx/h} g(x)g(y)dy\,dx\,. \tag{6.13}$$

We assume that

$$0 < h < k \tag{6.14}$$

and define $R(h,k)$ and l by the equations

$$V(h,k) = \frac{1}{2}G(h) - \frac{1}{2\pi}\arctan\frac{h}{k} + R(h,k)\,, \tag{6.15}$$

$$h^2 + k^2 = l^2\,, \qquad l > 0\,, \tag{6.16}$$

and obtain a divergent enveloping series for $R(h,k)$:

$$R(h,k) \propto \frac{h}{k}\frac{1}{2\pi l^2}e^{-l^2/2}\left[1 - \left(\frac{1}{k^2}+\frac{2}{l^2}\right) + \left(\frac{1\cdot 3}{k^4}+\frac{1\cdot 4}{k^2l^2}+\frac{2\cdot 4}{l^4}\right)\right.$$
$$\left. - \left(\frac{1\cdot 3\cdot 5}{k^6}+\frac{1\cdot 3\cdot 6}{k^4l^2}+\frac{1\cdot 4\cdot 6}{k^2l^4}+\frac{2\cdot 4\cdot 6}{l^6}\right) + \cdots\right]. \tag{6.17}$$

The general term of the series in brackets is of the dimension $-2n$ in k and l (or in h and k) and is itself a sum of $n + 1$ terms:

$$(-1)^n \sum_{\nu=0}^{n} \frac{1}{k^2}\frac{3}{k^2}\cdots\frac{2\nu-1}{k^2}\frac{2\nu+2}{l^2}\cdots\frac{2n-2}{l^2}\frac{2n}{l^2}\,. \tag{6.18}$$

7. Proof of the double inequality

In order to evaluate the double integrals L or M, defined by formulas (3) and (4) of the introduction, we transform the quadratic in x and x' into a sum of two squares by an appropriate substitution. There are various ways of doing so, and we first choose

$$y = \frac{x' - rx}{(1-r^2)^{1/2}}\,, \qquad dy = \frac{dx'}{(1-r^2)^{1/2}}\,, \tag{7.1}$$

which leads to

$$M(h,k;r) = \int_h^\infty g(x)\int_{\frac{k-rx}{(1-r^2)^{1/2}}}^\infty g(y)dy\,dx = \int_h^\infty g(x)\left[\frac{1}{2} - G\left(\frac{k-rx}{(1-r^2)^{1/2}}\right)\right]dx$$
$$= \int_h^\infty g(x)\left[1 - \left\{\frac{1}{2} - G\left(\frac{rx-k}{(1-r^2)^{1/2}}\right)\right\}\right]dx = \frac{1}{2} - G(h) - Q\,, \tag{7.2}$$

[9] *Biometrika*, vol. 33 (1943), pp. 59–72.

where

$$Q = \int_h^\infty g(x) \left\{ \frac{1}{2} - G\left(\frac{rx - k}{(1 - r^2)^{1/2}} \right) \right\} dx. \tag{7.3}$$

We used

$$G(-x) = -G(x)$$

which follows immediately from definition (2).

The integrand in (7.3) is visibly positive. Hence, and from (7.2), we deduce (6.2). In order to prove (6.3) also, we now seek an upper bound for Q.

We take into account the fact that the first term of the asymptotic series (6.12) is greater than the enveloped quantity on the left-hand side, and use conditions (6.1). We thus obtain from (7.3)

$$\begin{aligned} Q &< \int_h^\infty g(x) \frac{(1 - r^2)^{1/2}}{rx - k} g\left(\frac{rx - k}{(1 - r^2)^{1/2}} \right) dx \\ &< \frac{(1 - r^2)^{1/2}}{rh - k} \int_h^\infty g(x)\, g\left(\frac{rx - k}{(1 - r^2)^{1/2}} \right) dx \\ &= \frac{(1 - r^2)^{1/2}}{rh - k} \int_h^\infty g\left(\frac{x - rk}{(1 - r^2)^{1/2}} \right) g(k)\, dx \\ &= \frac{1 - r^2}{rh - k} \left[\frac{1}{2} - G\left(\frac{h - rk}{(1 - r^2)^{1/2}} \right) \right] g(k), \end{aligned} \tag{7.4}$$

which proves (6.3).

8. The expansion underlying the approximate formula

We now transform the quadratic in x and x' into a sum of two squares by a substitution different from the one used at the beginning of the foregoing section 7. For x and x' we write u and v respectively, and then we put

$$x = \frac{u + v}{[2(1 + r)]^{1/2}}, \qquad y = \frac{-u + v}{[2(1 - r)]^{1/2}}, \tag{8.1}$$

which implies

$$\frac{\partial(x,y)}{\partial(u,v)} = [1 - r^2]^{-1/2}.$$

The integral (3), extended over a rectangle, is changed into one extended over a parallelogram P with vertices (α,α'), (β,β'), (γ,γ'), and (δ,δ'); compare figure 1, drawn in the u,v-plane, with figure 2, in the x,y-plane, and observe that the four angles marked there are all equal. See also (6.5) and (6.6). Thus

$$L = \iint_P g(x)\, g(y)\, dx\, dy. \tag{8.2}$$

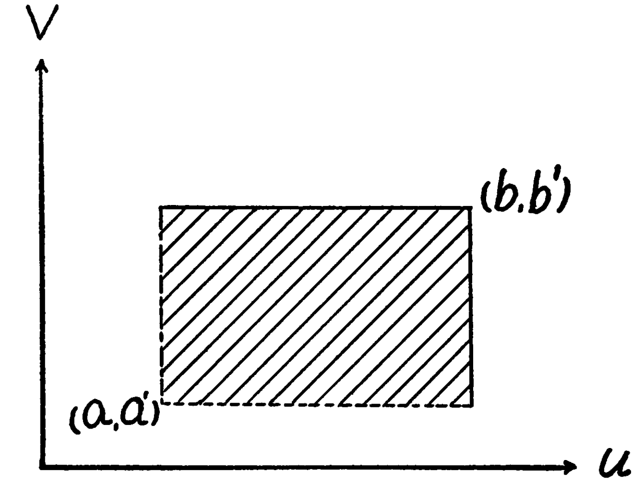

Figure 1

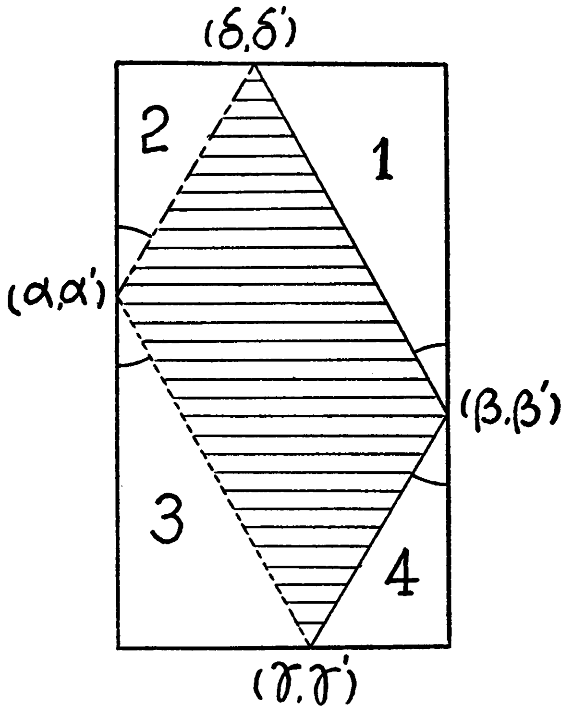

Figure 2

In order to evaluate the integral over the parallelogram P, we calculate the integral over the rectangle which contains P, as shown in figure 2, and subtract from it four integrals, denoted by Δ_1, Δ_2, Δ_3, and Δ_4, extended over the four triangles 1, 2, 3, and 4 shown in figure 2. That is,

$$(8.3)\quad L = \int_{\alpha}^{\beta}\int_{\gamma'}^{\delta'} g(x)\,g(y)\,dy\,dx - \Delta_1 - \Delta_2 - \Delta_3 - \Delta_4\,,$$

where

$$\Delta_2 = \int_{\alpha'}^{\delta'} g(y) \int_0^{\rho(y-\alpha')} g(\alpha + t)\, dt\, dy, \qquad \Delta_1 = \int_{\beta'}^{\delta'} g(y) \int_0^{\rho(y-\beta')} g(\beta - t)\, dt\, dy,$$

$$\Delta_3 = \int_{\gamma'}^{\alpha'} g(y) \int_0^{\rho(\alpha'-y)} g(\alpha + t)\, dt\, dy, \qquad \Delta_4 = \int_{\gamma'}^{\beta'} g(y) \int_0^{\rho(\beta'-y)} g(\beta - t)\, dt\, dy.$$

We have disposed here the four integrals Δ as the corresponding four triangles are disposed in figure 2. We put

$$x = \alpha + t \quad \text{in } \Delta_2 \text{ and } \Delta_3, \qquad x = \beta - t \quad \text{in } \Delta_1 \text{ and } \Delta_4.$$

We obtain finally, expanding the integrands of the four integrals Δ into powers of ρ (which is a small quantity when r is nearly one),

$$(8.4) \quad L(a,a';b,b';r) = [G(\beta) - G(\alpha)][G(\delta') - G(\gamma')]$$
$$- \sum_{n=0}^{\infty} \frac{g^{(n)}(\alpha)\, \rho^{n+1}}{(n+1)!} \int_{\alpha'}^{\delta'} (y - \alpha')^{n+1} g(y)\, dy$$
$$- \sum_{n=0}^{\infty} \frac{(-1)^n g^{(n)}(\beta)\, \rho^{n+1}}{(n+1)!} \int_{\beta'}^{\delta'} (y - \beta')^{n+1} g(y)\, dy$$
$$- \sum_{n=0}^{\infty} \frac{g^{(n)}(\alpha)\, \rho^{n+1}}{(n+1)!} \int_{\gamma'}^{\alpha'} (\alpha' - y)^{n+1} g(y)\, dy$$
$$- \sum_{n=0}^{\infty} \frac{(-1)^n g^{(n)}(\beta)\, \rho^{n+1}}{(n+1)!} \int_{\gamma'}^{\beta'} (\beta' - y)^{n+1} g(y)\, dy.$$

Retaining just the initial term, which corresponds to $n = 0$, of each of the four series, we obtain the approximate formula (6.7).

It is possible to estimate the remainders of the four series arising in (8.4) and, especially, the error of the approximate formula (6.7). The simplest estimate, however, is much too high for the purpose of computing. Roughly speaking, the approximate formula has not much chance to be good unless a differs but little from a' and b from b', but the simple (6.8) is particularly good. In practice, it is preferable to judge the goodness of approximation for a certain range of the parameters a,a', b,b' and r by comparing values given by (6.7) or (6.8) with values easily obtainable from the tables.

It may be mentioned that an expansion analogous to (8.4) and approximate formulas analogous to (6.7) and (6.8) can be obtained for M.

9. A lemma on enveloping series

In deriving the enveloping series (6.17) we meet with a certain situation which is much better understood when it is considered in full generality. Therefore we begin by explaining a general result.

LEMMA. *If*

$$a \propto b_1 - b_2 + b_3 - \cdots \tag{9.1}$$

and, for k $= 1, 2, 3, \cdots,$

$$b_k \propto c_{k1} - c_{k2} + c_{k3} - \cdots \tag{9.2}$$

then

$$a \propto c_{11} - (c_{12} + c_{21}) + (c_{13} + c_{22} + c_{31}) - \cdots. \tag{9.3}$$

The conclusion of our lemma asserts that a is enveloped by a series whose nth term is represented by

$$(-1)^{n-1}(c_{1n} + c_{2,n-1} + c_{3,n-2} + \cdots + c_{n1}).$$

The result can be restated without symbols: *If a number is enveloped by a series each term of which is enveloped by a series, then the resulting double series can be rearranged into another enveloping series for the same number by grouping terms in lines perpendicular to the main diagonal.*

In order to prove this lemma we have to make use of all the inequalities involved by the hypotheses (9.1) and (9.2), and the only difficulty is to group these inequalities suitably. There are two kinds of inequalities to prove: a must be shown to be less than certain partial sums of the series on the right-hand side of (9.3), and to be greater than certain other partial sums. It will suffice to consider the first kind of inequalities. Therefore we assume that n is *odd*, and derive from hypothesis (9.2) the following inequalities:

$$\begin{aligned}
b_1 &< c_{11} - c_{12} + c_{13} - \cdots + c_{1n} \\
b_2 &> \qquad c_{21} - c_{22} + \cdots - c_{2,n-1} \\
b_3 &< \qquad\qquad c_{31} - \cdots + c_{3,n-2} \\
&\cdots\cdots\cdots\cdots \\
b_n &< \qquad\qquad\qquad c_{n1}.
\end{aligned}$$

Multiply the second, fourth, sixth, $\cdots$ line by -1, add all these lines, and take into account the hypothesis (9.1). We obtain

$$\begin{aligned}
a &< b_1 - b_2 + b_3 - \cdots + b_n \\
&< c_{11} - (c_{12} + c_{21}) \\
&\quad + (c_{13} + c_{22} + c_{31}) \\
&\quad - \cdots\cdots \\
&\quad \cdots\cdots \\
&\quad + (c_{1n} + c_{2,n-1} + c_{3,n-2} + \cdots + c_{n1}),
\end{aligned}$$

and this proves our point. The case of even n can be treated similarly.

COROLLARY 1. *If*

$$A \propto a_1 - a_2 + a_3 - a_4 + \cdots \tag{9.4}$$

$$B \propto b_1 - b_2 + b_3 - b_4 + \cdots \tag{9.5}$$

and, moreover, $a_1 > 0$ *and* $B > 0$, *then*

$$(9.6) \qquad AB \propto a_1b_1 - (a_1b_2 + a_2b_1) + (a_1b_3 + a_2b_2 + a_3b_1) + \cdots .$$

It should be observed that the hypothesis (9.4) implies that $a_2, a_3, a_4, \cdots$ are positive; but we must postulate separately the positivity of a_1. The corollary is immediate. In fact, it follows directly from the hypothesis that

$$(9.7) \qquad AB \propto a_1B - a_2B + a_3B - \cdots$$

and that

$$(9.8) \qquad a_kB \propto a_kb_1 - a_kb_2 + a_kb_3 - \cdots ,$$

and so the hypotheses of the lemma are satisfied; (9.7) and (9.8) must be compared to (9.1) and (9.2), respectively.

COROLLARY 2. *If each of the positive quantities* A, B, $\cdots$ K, *and* L *is enveloped by a series, the product* AB $\cdots$ KL *is enveloped by the Cauchy product of these series.*

This proposition is slightly more general than a proposition recently obtained by J. V. Uspensky[10] and is derived by repeated application of corollary 1. We mention these corollaries because of their independent interest, but we shall not use them. We shall, however, use essentially the lemma in investigating the volume $V(h,k)$ defined by (6.13).

We dissect the first quadrant of the x,y-plane by producing the sides of the triangle over which the integral defining $V(h,k)$ is extended. We obtain

$$(9.9) \qquad V(h,k) = \left(\int_0^h \int_0^\infty - \int_0^\infty \int_{kx/h}^\infty + \int_h^\infty \int_{kx/h}^\infty\right) g(x)\, g(y)\, dy\, dx$$
$$= \frac{1}{2} G(h) - \frac{1}{2\pi} \arctan \frac{h}{k} + R(h,k) ;$$

in computing the middle integral we used transformation to polar coördinates in a well-known fashion. Thus

$$(9.10) \qquad R(h,k) = \int_h^\infty g(x) \int_{kx/h}^\infty g(y)\, dy\, dx .$$

Now, by (6.12),

$$(9.11) \qquad \int_{kx/h}^\infty g(y)\, dy \propto g(kx/h) \sum_{n=0}^\infty (-1)^n\, 1.3.5 \cdots (2n-1)\, (kx/h)^{-2n-1} .$$

[10] See *Mathematicae Notae*, vol. 4 (1944), pp. 1–10, especially pp. 2–4.

Observe that, with the abbreviation (6.16),

$$(9.12)\qquad g(x)\,g(kx/h) = (2\pi)^{-1}e^{-x^2l^2/(2h^2)}\,.$$

By (9.10), (9.11), and (9.12),

$$(9.13)\quad R(h,k) \propto (2\pi)^{-1}\sum_{n=0}^{\infty}(-1)^n 1.3.5\cdots(2n-1)(h/k)^{2n+1}\int_h^{\infty} x^{-2n-1}e^{-x^2l^2/(2h^2)}dx\,,$$

or

$$(9.14)\quad R(h,k) \propto h/(2\pi l)\sum_{n=0}^{\infty}(-1)^n\, 1.3.5\cdots(2n-1)\,(l/k)^{2n+1}I_{2n+1}(l)\,.$$

We introduce the abbreviation

$$(9.15)\qquad I_m(x) = \int_x^{\infty} t^{-m}\,e^{-t^2/2}\,dt\,.$$

Integrating by parts, we easily obtain

$$(9.16)\qquad I_m(x) = x^{-m-1}\,e^{-x^2/2} - (m+1)\,I_{m+2}(x)\,,$$

and repeating this we obtain the (divergent) enveloping series

$$(9.17)\quad I_m(x) \propto e^{-x^2/2}[x^{-m-1} - (m+1)x^{-m-3} + (m+1)\,(m+3)x^{-m-5} - \cdots]\,.$$

Therefore

$$(9.18)\quad 1.3.5.\cdots(2n-1)\,(l/k)^{2n+1}\,I_{2n+1}(l)$$
$$\propto (kl)^{-1}\,e^{-l^2/2}\left[\frac{1.3\cdots(2n-1)}{k^{2n}} - \frac{1.3\cdots(2n-1)(2n+2)}{k^{2n}l^2} + \frac{1.3\cdots(2n-1)\,(2n+2)\,(2n+4)}{k^{2n}l^4} - \cdots\right].$$

Thus $R(h,k)$ is enveloped by the series (9.14), each term of which is enveloped, as (9.18) shows, by a series. Applying our lemma, we obtain (6.17).[11]

[11] [Added November 19, 1945.] After this paper was written, a discussion with Miss Madeline Johnsen led to recognizing that the series (6.17) coincides essentially with one considered by W. F. Sheppard, "On the calculation of the double integral expressing normal correlation," *Trans. Cambridge Philos. Soc.*, vol. 19 (1904), pp. 23–68; see p. 37, formula (68). It can be observed, however, that: (*a*) the series appears here in a simpler form; and (*b*) Sheppard does not even mention the property proved here, that the series is enveloping. Miss Johnsen proves this property in a quite different way in her dissertation "Approximate evaluation of double probability integrals" deposited in the Library of Stanford University. See *Abstract of Dissertations, Stanford University*, vol. 21 (1945–46), pp. 113–116.