# ON THE REPRESENTATION
# OF A CERTAIN FUNDAMENTAL LAW OF PROBABILITY*

BY

H. L. RIETZ

1. **Introduction.** In his *Théorie Analytique des Probabilités*,[†] Laplace makes use of a function which we may write in the form

$$(1) \quad f(x) = \frac{1}{a^n (n-1)!} \left[ x^{n-1} - \binom{n}{1} (x-a)^{n-1} + \cdots + (-1)^n (x-na)^{n-1} \right]$$

in which each expression $(x-ma)^{n-1}$, $m = 0, 1, 2, \cdots, n$, is assigned the value zero if the number $x-ma$ is not positive.

With $f(x)$ thus defined, Laplace found that

$$f(x)\, dx$$

gives, to within infinitesimals of higher order, the probability that the sum of $n$ elements each taken at random from a given range 0 to $a$ ($a > 0$) of uniform distribution, will fall into the interval $x$ to $x + dx$.

Laplace applied the above formula to the historic problem[‡] of finding the probability that the inclinations of the orbits of the ten planets besides the earth known at the beginning of the year 1801 do not constitute a random distribution.

When Laplace proceeded to apply his theory to the orbits of all comets[§] known at the end of the year 1811, the number $n$ in formula (1) was so large as to render the formula impracticable for numerical computation. Laplace[||] used the form

$$(2) \quad \frac{(n + r\sqrt{n})^{n-1} - \binom{n}{1} (n + r\sqrt{n} - 2)^{n-1} + \binom{n}{2} (n + r\sqrt{n} - 4)^{n-1} - \cdots}{2^n (n-1)!},$$

---

where $(n + r\sqrt{n})/2 = x/a$, in place of $\frac{1}{2}f(x)$.  He used[*]

(3)
$$\sqrt{\frac{3}{2n\pi}}\, e^{-3r^2/2}\left[1 - \frac{3}{20n}(1 - 6r^2 + 3r^4) + \cdots\right]$$

as an approximation to (2) and the integral of (3) as an approximation
to the corresponding integral of (2).

It was recognized by Cauchy that the methods used by Laplace in ob-
taining the approximations were lacking in rigor.  In fact, Cauchy[†] put
the results on a much more secure basis in a memoir published in 1841.
Todhunter[‡] described this memoir as "very laborious and difficult, so that
this portion of the théorie des probabiltés remains in an unsatisfactory state."

It is the main object of the present paper to put this portion of the
theory of probability into a more satisfactory state by developing an ap-
proximation to $f(x)$ by means of a Gram-Charlier series.  It results that
the approximation used by Laplace is given by the first two non-vanishing
terms of this series, and that closer and closer approximations in the sense
of the theory of least squares are obtained by the use of each additional
term of the series.  The approximation obtained by retaining three non-
vanishing terms of the Gram-Charlier series agrees exactly with the corre-
sponding terms of the Cauchy series, but the derivation of Cauchy does
not seem to imply that his approximation is best in the sense of a least
squares criterion.

The main difficulty and much of the interest in our method of finding the
approximation consists in obtaining remarkably simple expressions for the
moments of area under the theoretical frequency curve.  Certain auxiliary
theorems proved in this connection seem to be new and to be of interest
in combinatory analysis.

2. **Symmetry of the curve $y = f(x)$ about the line $x = na/2$.**
It will be convenient later in this paper (p. 204) to use the property that
the curve $y = f(x)$ is symmetrical about the line $x = na/2$.  To prove
the symmetry in question, we may first assume that the probability that
the sum of the distances of $n$ random points on the line $AB$ from the
point $A$ will fall into an assigned interval $x_1$ to $x_1 + dx$

$$\begin{array}{ccc} A & & B \\ \hline 0 & & a \end{array}$$

[*] Loc. cit., p. 173.

[†] *Mémoire sur divers formules relatives à la théorie des intégrales définies et sur la
conversion des différences finies des puissances en intégrales de cette espèce*, Journal de
L'Ecole Polytechnique, Cahier 28, Tome 17 (1841), pp. 147–248.

[‡] Todhunter, *History of Probability*, 1865, p. 527.

is the same as the probability that the sum of the distances of $B$ from the $n$ points will fall into the same interval. Next, it is obvious that a sum of distances of such $n$ points from $A$ falls into the interval $x_1$ to $x_1 + dx$ when and only when the corresponding sum of distances from the points to $B$ falls into the interval $na - x_1 - dx$ to $na - x_1$. Hence, the probability that a sum of distances of the $n$ points from $A$ will fall into the interval $x_1$ to $x_1 + dx$ is equal to the probability that the sum will fall into the interval $na - x_1 - dx$ to $na - x_1$, and the symmetry of the probability curve $y = f(x)$ about $x = na/2$ is established.

**3. On the convergence of the Gram-Charlier series representing $f(x)$.** The function $y = f(x)$ $(n \geq 2)$ may obviously be regarded as the sum of a set of rational integral functions with junction points at $x = a, 2a, \cdots, (n-1)a$ and with end points at $x = 0$ and $x = na$. Hence, $y = f(x)$ and its successive derivatives are continuous from $x = 0$ to $x = na$ unless it be at junction points. It is easily shown* that $y = f(x) = 0$ when $x \geq na$, and we shall find it useful for our purposes to use this extended interval of definition of the function. Furthermore, we shall find it convenient to extend the interval of $y = f(x)$ from $x = 0$ to $-\infty$ by making $y = f(x) = 0$ throughout this extension. Then the end points $x = 0$ and $x = na$ of the original interval 0 to $na$ become junction points.

It follows that $f(x)$ is continuous at any junction point $x = pa$ because the additional term of (1), added for the interval $pa < x < (p+1)a$, has the limit zero as $x \to pa$.

The first derivative of (1),

$$(4) \quad \frac{dy}{dx} = \frac{1}{a^n (n-2)!} \left\{ x^{n-2} - \binom{n}{1} (x-a)^{n-2} + \binom{n}{2} (x-2a)^{n-2} - \cdots \right. $$
$$\left. + (-1)^n \binom{n}{n} (x-na)^{n-2} \right\},$$

is likewise continuous at the points $x = 0, 2a, \cdots, na$, except when $n = 2$. When $n = 2$ and $x = a$, we have

$$\frac{dy}{dx} = \frac{1}{a^2} \quad \text{or} \quad -\frac{1}{a^2}$$

according as $x \to a$ from below or above $a$.

---

* Cf. Chrystal's *Algebra*, Part 2, 1889, p. 210.

The second derivative of (1),

$$\frac{d^2 y}{dx^2} = \frac{1}{a^n (n-3)!} \left\{ x^{n-3} - \binom{n}{1}(x-a)^{n-3} + \binom{n}{2}(x-2a)^{n-3} - \cdots \right.$$
$$\left. + (-1)^n \binom{n}{n}(x-na)^{n-3} \right\},$$

(5)

is clearly continuous when $n > 3$. It is continuous when $n = 3$ except at $x = 0$, $a$, $2a$, and $3a$.

For example, when $n = 3$ and $x = 2a$, we have

$$\frac{d^2 y}{dx^2} = -\frac{2}{a^3} \text{ or } +\frac{1}{a^3}$$

according as $x \to 2a$ from below or above $2a$.

The examination of the function for continuity could be extended in an obvious manner to higher derivatives, but we shall be interested mainly in the continuity of $f(x)$ and its first and second derivatives.

The theory[*] of the Gram-Charlier representation of an arbitrary frequency function may be based on the theorem[†] of W. Myller-Lebedeff proved by the use of integral equations, that any function $g(z)$ which together with its first and second derivatives is finite and continuous from $-\infty$ to $+\infty$ and for which

$$\lim_{z = \pm \infty} z^3 g(z) = 0$$

can be represented by an absolutely uniformly convergent infinite series of the form

(6)             $c_0 \Phi_0(z) + c_1 \Phi_1(z) + c_2 \Phi_2(z) + \cdots + c_i \Phi_i(z) + \cdots,$

where $\Phi_i(z)$ is the product of the Gaussian probability function

$$\Phi_0(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

and the $i$th order polynomial of Hermite $H_i(z)$.

---

[*] J. P. Gram, Journal für Mathematik, vol. 94 (1883), pp. 41-73; also dissertation, Copenhagen, 1879.

C. L. V. Charlier, *Über die Darstellung willkürlicher Funktionen*, Arkiv for Matematik, Astronomi och Fysik, vol. 2, No. 20 (1905-1906), pp. 1-33. *Vorlesungen über die Grundzüge der mathematischen Statistik*, 1920, pp. 67-78.

Cf. H. Bruns, *Wahrscheinlichkeitsrechnung und Kollektivmasslehre*, 1906, pp. 115-125.

Cf. Mises, Jahresbericht der deutschen Mathematiker-Vereinigung, vol. 21 (1912), pp. 9-20.

[†] W. Myller-Lebedeff, Mathematische Annalen, vol. 64 (1907), p. 400.

We have shown that the conditions of continuity are satisfied when $n > 3$. Obviously the condition

$$\lim_{z=\pm\infty} x^3 f(x) = 0$$

is also satisfied.

In this representation by (6) the $\Phi$-functions and the $H$-functions form a biorthogonal system. Thus

$$\int_{-\infty}^{+\infty} \Phi_m(z) H_n(z) \, dz = 0 \quad \text{if} \quad m \neq n.$$

The representation of $f(x)$ by such a convergent infinite series naturally suggests that the first few terms of (6) might give a valuable approximation, but the representation is particularly appropriate because it can be shown to be best in the sense of the theory of least squares.

**4. Change of origin and unit.** It is well known* that by both choosing the origin of coördinates so that the $x$-coördinate of the centroid of area bounded by the graph of the function to be represented and the $x$-axis is equal to zero, and choosing the standard deviation (radius of gyration), $\sigma$, as the unit of measurement, we have

$$c_1 = 0 \quad \text{and} \quad c_2 = 0.$$

We accordingly make the transformation

$$\frac{x - \dfrac{n a}{2}}{\sigma} = z$$

and let

(7)                    $$y = f(x) = g(z).$$

Moreover, we shall see presently that any coefficient $c_i$, when $i$ is odd, vanishes because of the symmetry of $y = g(z)$ about $z = 0$ which was proved in § 2. Thus

(8)                    $$c_i = \frac{(-1)^i}{i!} \int_{-\infty}^{+\infty} g(z) H_i(z) \, dz,$$

---

* H. Bruns, loc. cit., pp. 118–119.
  C. V. L. Charlier, loc. cit., pp. 62-67.

where

$$(9) \quad H_i(z) = z^i - \frac{i(i-1)}{2} z^{i-2} + \frac{i(i-1)(i-2)(i-3)}{2 \cdot 4} z^{i-4} - \cdots$$

is the Hermite polynomial of order $i$.

The determination of $c_i$ from (8) would, when $i$ is odd, clearly involve merely the sum of a set of odd order moments of area about an axis of symmetry so that each moment is zero. Hence, $c_i = 0$ when $i$ is odd.

We may then write in place of (6)

$$(10) \qquad g(z) = c_0 \, \Phi_0(z) + c_4 \, \Phi_4(z) + c_6 \, \Phi_6(z) + \cdots.$$

Let us examine the finite series

$$(11) \qquad U(z) = c_0 \, \Phi_0(z) + c_4 \, \Phi_4(z) + \cdots + c_{2l} \, \Phi_{2l}(z)$$

as an approximate representation of $g(z)$.

When the coefficients $c_0, c_2, \cdots, c_{2l}$ in (11) are given the values in (8) obtained by the use of the biorthogonal property of the $\Phi$ and $H$-functions, it is known that we have the best approximation of $U(z)$ to $g(z)$ in the sense that a certain least square criterion*

$$(12) \qquad I = \int_{-\infty}^{+\infty} \frac{1}{\Phi_0(z)} [g(z) - U(z)]^2 \, dz$$

is a minimum.

Moreover, it is known from a general theory of Gram that the approximation of $U(z)$ to $g(z)$ becomes closer[†] and closer with each increase of $l$ in (11).

In the use of the least squares criterion (12), a question naturally arises as to the propriety of weighting squares of deviations

$$\delta^2 = [g(z) - U(z)]^2$$

with the reciprocal

$$\frac{1}{\Phi_0(z)} = \sqrt{2\pi} \, e^{z^2/2}$$

of the Gaussian probability function. Gram uses this weighting without commenting on its propriety so far as I have been able to find. One fairly

---

\* Gram, loc. cit., p. 54; cf. Mises, loc. cit., p. 20.

† Loc. cit., pp. 47–55.

obvious point in support of such weighting is its algebraic convenience. Another point suggestive of the given weighting is found in the fact that the reciprocal of $\Phi_0(z)$ as a weight tends to weaken* the effect of the square of $\Phi_0(z)$ which occurs as a factor in $[g(z) - U(z)]^2$. Furthermore, it seems natural to have regard for algebraic convenience in this connection because practical statisticians have pointed out the fact that different sets of weights frequently lead to almost identical results. While a sort of rationale thus underlies the scheme of weighting, there remains an element of arbitrariness about it in the sense that it is not necessarily better than some other weighting which might be proposed.

5. **Coefficient of the $\Phi$-series expressed in moments.** We shall give special attention to the approximation by the use of the first three terms of (11) because we obtain the approximation which Laplace used from the first two terms, and a closer approximation in the sense of the theory of least squares by retaining three terms.

The coefficients $c_0$, $c_4$, $c_6$ are now easily expressed in terms of the area $A = 1$ under the curve $y = f(x)$ and the moments of this area about the axis $x = na/2$ through the centroid. Let $\mu_q$ be the $q$th moment of area about this axis. Then

$$(13) \qquad \mu_q = \int_0^{na} \left(x - \frac{na}{2}\right)^q f(x)\, dx = \sigma^{q+1} \int_{-na/2\sigma}^{na/2\sigma} z^q g(z)\, dz,$$

from the definition of $g(z)$ in (7). Then by means of (8) and (13), we obtain

$$(14) \qquad c_0 = \frac{1}{\sigma},$$

$$(15) \qquad c_4 = \frac{1}{4!\,\sigma}\left[\frac{\mu_4}{\sigma^4} - 3\right],$$

$$(16) \qquad c_6 = \frac{1}{6!\,\sigma}\left[\frac{\mu_6}{\sigma^6} - 15\frac{\mu_4}{\sigma^4} + 30\right].$$

By recalling that $\sigma = \sqrt{\mu_2}$, we may say that the difficulties of our problem have now been reduced to that of expressing the moments $\mu_2$, $\mu_4$, and $\mu_6$ in simple form. In this process we shall derive some remarkably simple results from somewhat complicated combinatory forms.

6. **Area and moments of area under** $y = f(x)$. From the definition of $f(x)$ it is fairly obvious that the area, $A$, is 1, but we present a veri-

---

* Suggested by Professor E. L. Dodd.

fication of this fact because we lead up in this way to the method by which we have found the moments of area. We should keep especially in mind in this connection that each parenthesis $(x - pa)$ of

$$(17) \quad y = \frac{1}{a^n (n-1)!} \left[ x^{n-1} - \binom{n}{1}(x-a)^{n-1} + \binom{n}{2}(x-2a)^{n-1} - \cdots \right.$$
$$\left. + (-1)^n \binom{n}{n}(x-na)^{n-1} \right]$$

is assigned the value zero when $x - pa$ is not positive.

The area would be found by integrating the first term of the right hand member of (17) from 0 to $a$, the first two terms from $a$ to $2a$, the first three terms from $2a$ to $3a$ and so on. But this process is equivalent to integrating the first term from 0 to $na$, the second from $a$ to $na$, the third from $2a$ to $na$ and so on to all the terms. Thus,

$$a^n(n-1)! A = \int_0^{na} x^{n-1} dx - \binom{n}{1} \int_a^{na} (x-a)^{n-1} dx + \binom{n}{2} \int_{2a}^{na} (x-2a)^{n-1} dx - \cdots$$
$$+ (-1)^{n-1} \binom{n}{n-1} \int_{(n-1)a}^{na} [x - (n-1)a]^{n-1} dx$$

$$(18) \qquad = \frac{a^n}{n} \left[ n^n - \binom{n}{1}(n-1)^n + \binom{n}{2}(n-2)^n - \cdots + (-1)^{n-1} \binom{n}{n-1} \right]$$

$$(19) \qquad = a^n(n-1)!.*$$

Hence,

$$(20) \qquad\qquad\qquad\qquad A = 1.$$

*Moments of odd order.* The moments of odd order of the area about its centroid vertical $x = na/2$ are each equal to zero, because of the symmetry about this axis. Hence we have

$$\mu_1 = \mu_3 = \mu_5 = \cdots = 0.$$

*Moments about the axis $x = na$.* We shall first find moments about the axis $x = na$, then after considerable simplification, we shall express the moments about the centroidal axis $x = na/2$ in terms of the moments about $x = na$. We let $\mu_q'$ represent the $q$th moment of the area about the axis $x = na$.

---

* Czuber, *Wahrscheinlichkeitsrechnung,* vol. I, third edition, 1914, p. 66.

(a) *Second moments.* In this case we have

$$a^n(n-1)!\,\mu_2' = \int_0^{na} x^{n-1}(x-na)^2\,dx - \binom{n}{1}\int_a^{na}(x-a)^{n-1}(x-na)^2\,dx + \cdots$$

$$+ (-1)^{n-1}\binom{n}{n-1}\int_{na-a}^{na}(x-na+a)^{n-1}(x-na)^2\,dx$$

$$= \int_0^{na} x^{n-1}(x-na)^2\,dx - \binom{n}{1}\int_a^{na}(x-a)^{n-1}[x-a-(n-1)a]^2\,dx + \cdots$$

$$+ (-1)^{n-1}\binom{n}{n-1}\int_{na-a}^{na}(x-na+a)^{n-1}(x-na+a-a)^2\,dx$$

$$= \left(\frac{1}{n+2}-\frac{2}{n+1}+\frac{1}{n}\right)a^{n+2}\left[n^{n+2}-\binom{n}{1}(n-1)^{n+2}+\binom{n}{2}(n-2)^{n+2}-\cdots\right.$$

$$\left.+ (-1)^{n-1}\binom{n}{n-1}\right]$$

$$= \frac{2\,a^{n+2}}{n(n+1)(n+2)}\left[n^{n+2}-\binom{n}{1}(n-1)^{n+2}+\binom{n}{2}(n-2)^{n+2}-\cdots\right.$$

$$\left.+ (-1)^{n-1}\binom{n}{n-1}\right],$$

(21)
$$\mu_2' = \frac{2\,a^2}{(n+2)!}\left[n^{n+2}-\binom{n}{1}(n-1)^{n+2}+\binom{n}{2}(n-2)^{n+2}-\cdots\right.$$

$$\left.+ (-1)^{n-1}\binom{n}{n-1}\right],$$

where the terms $(n-p)$ are assigned the value zero whenever $n-p$ is not positive.

(b) *Fourth moments.* In this case,

$$a^n(n-1)!\,\mu_4' = \int_0^{na} x^{n-1}(x-na)^4\,dx - \binom{n}{1}\int_a^{na}(x-a)^{n-1}(x-na)^4\,dx + \cdots$$

$$+ (-1)^{n-1}\binom{n}{n-1}\int_{na-a}^{na}(x-na+a)^{n-1}(x-na)^4\,dx$$

$$= \int_0^{na} x^{n-1}(x-na)^4\,dx - \binom{n}{1}\int_a^{na}(x-a)^{n-1}[x-a-na+a]^4 + \cdots$$

$$+ (-1)^{n-1}\binom{n}{n-1}\int_{na-a}^{na}(x-na+a)^{n-1}(x-na+a-a)^4\,dx$$

$$= \left( \frac{1}{n+4} - \frac{4}{n+3} + \frac{6}{n+2} - \frac{4}{n+1} + \frac{1}{n} \right) a^{n+4} \left[ n^{n+4} - \binom{n}{1}(n-1)^{n+4} \right.$$

$$\left. + \binom{n}{2}(n-2)^{n+4} - \cdots + (-1)^{n-1}\binom{n}{n-1} \right]$$

$$= \frac{a^{n+4}\,4!}{n(n+1)(n+2)(n+3)(n+4)} \left[ n^{n+4} - \binom{n}{1}(n-1)^{n+4} \right.$$

$$\left. + \binom{n}{2}(n-2)^{n+4} - \cdots + (-1)^{n-1}\binom{n}{n-1} \right]$$

$$\mu_4' = \frac{a^4\,4!}{(n+4)!} \left[ n^{n+4} - \binom{n}{1}(n-1)^{n+4} + \binom{n}{2}(n-2)^{n+4} - \cdots \right.$$

(22)
$$\left. + (-1)^{n-1}\binom{n}{n-1} \right].$$

(c) *General formula for moments of order m (m any positive integer).*
By a process similar to that employed for second and fourth moments,
we find the $m$th moment about $x = na$ to be

$$\mu_m' = (-1)^m \frac{a^m\,m!}{(n+m)!} \left[ n^{n+m} - \binom{n}{1}(n-1)^{n+m} + \binom{n}{2}(n-2)^{n+m} - \cdots \right.$$

(23)
$$\left. + (-1)^{n-1}\binom{n}{n-1} \right],$$

where $m$ is any positive integer.

7. **On the reduction of certain combinatory forms.** The moments
of area about the axis $x = na$ obtained in § 6 involve combinatory
expressions of the form

$$F(n, t) = \frac{t!}{(n+t)!} \left[ n^{n+t} - \binom{n}{1}(n-1)^{n+t} + \binom{n}{2}(n-2)^{n+t} - \cdots \right.$$

(24)
$$\left. + (-1)^{n-1}\binom{n}{n-1} \right],$$

where $t$ is an integer $\geq 1$.

Although it is well known that $F(n, 0) = 1$, and that the expression in
brackets in the right hand member of (24) is equal* to zero when $t$ is
a negative integer $(n + t > 0)$, no simple values seem to have been found
for $F(n, t)$ when $t \geq 1$, so far as the writer has been able to learn.

In what follows, we obtain simple polynomials in $n$ for $F(n, t)$ when
$t = 1, 2, 3, \cdots, 6$, and find (§ 9) that the results lead to very important
simplifications in the first six moments given in § 5. Similar results could

---

* Picquet, Journal de Mathématiques Spéciales, ser. 3, vol. 2 (1888), pp. 150, 172,
196.  E. Netto, Lehrbuch der Combinatorik, 1901, p. 37.

be obtained for somewhat larger values of $t$ if they were needed.  In fact the writer has derived the results for $t = 7$ and $t = 8$.  The following theorems giving $F(n, t)$ for $t = 1, 2, 3, \cdots, 6$, in the form of polynomials will be proved:

$$(25) \qquad F(n, 1) = \frac{n}{2};$$

$$(26) \qquad F(n, 2) = \frac{n(3n+1)}{12};$$

$$(27) \qquad F(n, 3) = \frac{n^2(n+1)}{8};$$

$$(28) \qquad F(n, 4) = n\left(\frac{n^3}{16} + \frac{n^2}{8} + \frac{n}{48} - \frac{1}{120}\right);$$

$$(29) \qquad F(n, 5) = \frac{n^2}{8}\left(\frac{n^3}{4} + \frac{5n^2}{6} + \frac{5n}{12} - \frac{1}{6}\right);$$

$$(30) \qquad F(n, 6) = \frac{n}{4}\left(\frac{n^5}{16} + \frac{5n^4}{16} + \frac{5n^3}{16} - \frac{13n^2}{144} - \frac{n}{24} + \frac{1}{63}\right).$$

It may be observed from (23) and (24) that the left hand members of (25), (27) and (29) are involved in odd order moments of area about the axis $x = na$.

Thus for the first moment about the axis $x = na$ we obtain from (23)

$$\mu_1' = -\frac{a}{(n+1)!}\left[n^{n+1} - \binom{n}{1}(n-1)^{n+1} + \binom{n}{2}(n-2)^{n+1} - \cdots \right.$$
$$\left. + (-1)^{n-1}\binom{n}{n-1}\right].$$

Then using (24) with $t = 1$,

$$(31) \qquad\qquad\qquad \mu_1' = -aF(n, 1).$$

But

$$(32) \qquad\qquad\qquad \mu_1' = -\frac{na}{2},$$

since the first moment coefficient $\mu_1'$ is simply the distance from the axis of moments to the centroid.

From (31) and (32), we have

$$(33) \qquad\qquad\qquad F(n, 1) = \frac{n}{2}.$$

Before presenting the proofs of the equalities (26) to (30), it may be of interest to explain what suggested the values given for $F(n, t)$ in each case. The value of $F(n, 2)$ was suggested by experimenting with the special cases $n = 1, 2, \cdots, 6$. The values of $F(n, 3)$ and $F(n, 5)$ were suggested by equating third and fifth order moments as given by (23) to the corresponding values found by transforming the odd order moments about $x = na/2$ to the parallel axis $x = na$. The values of $F(n, 4)$ and $F(n, 6)$ were suggested by finding the fourth and sixth moments about $x = na/2$ of the area under the Gaussian probability curve which would fit $y = f(x)$ best, and equating these expressions for the moments to the values of the corresponding moments $\mu_4$ and $\mu_6$ obtained from $\mu_4'$ and $\mu_6'$ given in § 6, and then experimenting with corrections until a correctional formula was found which could be proved to hold for any integral value of $n$.

We shall now prove (26) to (30) by mathematical induction. First, verify the relations for small values of $n$, say for $n = 1, 2, 3$.

In the proof, we make much use of the relation*

$$(34) \qquad (k + t)F(k, t) = k[tF(k, t-1) + F(k-1, t)],$$

which is easily established by substitution from the definition of $F(n, t)$ given in (24), or by the use of a theorem† of combinatory analysis to which (34) is closely related.

By means of (34) we can now prove that each of the equalities (26) to (30) is true for a positive integer $n = k$ if it is true for $n = k - 1$.

(a) First consider formula (26). From (34),

$$(k + 2)F(k, 2) = k\{2F(k, 1) + F(k-1, 2)\}$$
$$= k\left\{k + \frac{k-1}{12}[3(k-1)+1]\right\},$$

since from (33), $F(k, 1) = k/2$, and since we are assuming the equality (26) for $n = k - 1$. Solving for $F(k, 2)$, we have

$$F(k, 2) = \frac{k(3k+1)}{12},$$

which was to be proved.

---

* Cf. J. Worpitzky, *Studien über die Bernoullischen und Eulerschen Zahlen*, Journal für Mathematik, vol. 94 (1883), p. 210.

† E. Netto, *Lehrbuch der Combinatorik,* 1901, p. 169.

($b$) Formula (27).　From (34),

$$(k+3)F(k, 3) = k[3F(k, 2) + F(k-1, 3)]$$
$$= k\left[\frac{k(3k+1)}{4} + \frac{k(k-1)^2}{8}\right] = \frac{k^2}{8}(k+1)(k+3),$$

since $F(k, 2) = k(3k+1)/12$, and we assume (27) for $n = k-1$. Hence,

$$F(k, 3) = \frac{k^2}{8}(k+1),$$

which was to be proved.

　($c$) Formula (28).　From (34),

$$(k+4)F(k, 4) = k[4F(k, 3) + F(k-1, 4)]$$
$$= k\left[\frac{k^3(k+1)}{2} + \frac{(k-1)^4}{16} + \frac{(k-1)^3}{8} + \frac{(k-1)^2}{48} - \frac{(k-1)}{120}\right]$$
$$= k(k+4)\left(\frac{k^3}{16} + \frac{k^2}{8} + \frac{k}{48} - \frac{1}{120}\right),$$

since $F(k, 3) = k^2(k+1)/8$, and we assume (28) for $n = k-1$. Hence,

$$F(k, 4) = k\left(\frac{k^3}{16} + \frac{k^2}{8} + \frac{k}{48} - \frac{1}{120}\right),$$

which was to be proved.

　($d$) Formula (29).　From (34),

$$(k+5)F(k, 5) = k[5F(k, 4) + F(k-1, 5)]$$
$$= k\left[\frac{5k^4}{16} + \frac{5k^3}{8} + \frac{5k^2}{48} - \frac{k}{24} + \frac{(k-1)^5}{32} + \frac{5(k-1)^4}{48}\right.$$
$$\left. + \frac{5(k-1)^3}{96} - \frac{(k-1)^2}{48}\right]$$
$$= \frac{k^2}{8}(k+5)\left(\frac{k^3}{4} + \frac{5k^2}{6} + \frac{5k}{12} - \frac{1}{6}\right).$$

Hence,

$$F(k, 5) = \frac{k^2}{8}\left(\frac{k^3}{4} + \frac{5k^2}{6} + \frac{5k}{12} - \frac{1}{6}\right),$$

which was to be proved.

($e$) Formula (30).   From (34),

$$(k+6)F(k, 6) = k[6F(k, 5)+F(k-1, 6)]$$

$$= k\left[\frac{6k^5}{32} + \frac{5k^4}{8} + \frac{5k^3}{16} - \frac{k^2}{8} + \frac{(k-1)^6}{64} + \frac{5(k-1)^5}{64}\right.$$

$$\left. + \frac{5(k-1)^4}{64} - \frac{13(k-1)^3}{576} - \frac{(k-1)^2}{96} + \frac{k-1}{252}\right]$$

$$= \frac{k}{4}(k+6)\left(\frac{k^5}{16} + \frac{5k^4}{16} + \frac{5k^3}{16} - \frac{13k^2}{144} - \frac{k}{24} + \frac{1}{63}\right).$$

Hence,

$$F(k, 6) = \frac{k}{4}\left(\frac{k^5}{16} + \frac{5k^4}{16} + \frac{5k^3}{16} - \frac{13k^2}{144} - \frac{k}{24} + \frac{1}{63}\right),$$

which was to be proved.

8. **Simplification of moments.** By the use of auxiliary theorems proved in § 7, we are now able to write the moments of area about the axis $x = na$ given in § 6 in greatly simplified forms.  Thus, by substitution from (26), (28) and (33) in (23) we obtain

$$(35) \qquad \mu_2' = \frac{n}{12}(3n+1)a^2,$$

$$(36) \qquad \mu_4' = \frac{n}{8}\left(\frac{n^3}{2} + n^2 + \frac{n}{6} - \frac{1}{15}\right)a^4,$$

$$(37) \qquad \mu_6' = \frac{n}{4}\left(\frac{n^5}{16} + \frac{5n^4}{16} + \frac{5n^3}{16} - \frac{13n^2}{144} - \frac{n}{24} + \frac{1}{63}\right)a^6.$$

9. **Even order moments about the axis $x = na/2$ through the centroid of area.**  The moments $\mu_2$, $\mu_4$, $\mu_6$ about the parallel axis through the centroid are given in terms of $\mu_2'$, $\mu_4'$, $\mu_6'$, and the distance $d = na/2$ between the parallel axes, by the well known relations

$$(38) \quad \mu_2 = \mu_2' - d^2,$$

$$(39) \quad \mu_4 = \mu_4' + 4\mu_3 d - 6\mu_2 d^2 + 4\mu_1 d^3 - d^4,$$

$$(40) \quad \mu_6 = \mu_6' + 6\mu_5 d - 15\mu_4 d^2 + 20\mu_3 d^3 - 15\mu_2 d^4 + 6\mu_1 d^5 - d^6.$$

Substituting in (38), (39) and (40) from (35), (36) and (37), remembering that odd order moments $\mu_1$, $\mu_3$, $\mu_5$ are each zero, we have

$$(41) \qquad \mu_2 = \frac{a^2 n}{12},$$

$$(42) \qquad \mu_4 = \frac{a^4 n}{24} \left( \frac{n}{2} - \frac{1}{5} \right),$$

$$(43) \qquad \mu_6 = \frac{a^6 n}{12} \left( \frac{5 n^2}{48} - \frac{n}{8} + \frac{1}{21} \right).$$

**10. The coefficients of the $\Phi(z)$-series.** By substitution for the moments $\mu_2$, $\mu_4$, and $\mu_6$ in (15) and (16) their simple values given in (41), (42) and (43), we have

$$(44) \qquad c_4 = - \frac{1}{20 n \sigma},$$

and

$$(45) \qquad c_6 = + \frac{1}{105 n^2 \sigma},$$

which seem to be remarkably simple results.

Hence, we have, from (10), (14), (44) and (45), for the approximate representation of our frequency function

$$(46) \qquad y = \frac{1}{\sigma} \left[ \varphi_0(z) - \frac{1}{20 n} \varphi_4(z) + \frac{1}{105 n^2} \varphi_6(z) \right],$$

where

$$\varphi_0(z) = \frac{1}{\sqrt{2 \pi}} e^{-z^2/2},$$

$$z = \frac{x - \dfrac{n a}{2}}{\sigma},$$

and

$$\sigma^2 = \frac{a^2 n}{12}.$$

The equation (46) may also be written in the form

$$y = \frac{1}{a} \sqrt{\frac{6}{n \pi}} e^{-z^2/2} \left[ 1 - \frac{1}{20 n} (z^4 - 6 z^2 + 3) \right.$$
$$(47) \qquad\qquad \left. + \frac{1}{105 n^2} (z^6 - 15 z^4 + 45 z^2 - 15) \right].$$

The first term of (46) is the Gaussian probability function. The sum of the first two terms divided by 2 is identical with the approximation which Laplace used for formula (2) as may be shown by making the transformations

$$z = \frac{x - \dfrac{na}{2}}{\sigma}, \qquad x = \frac{n + r\sqrt{n}}{2}, \qquad a = 1.$$

## 11. Numerical checks on approximations.
Although it is known (§ 4) that our approximate representation of $y = f(x)$ is improved by the inclusion of each additional term of (11) in the sense of minimizing a certain least squares criterion, we find by numerical computation that the approximation is not in all cases improved at every point by the inclusion of an additional term. In order to emphasize this fact, and to gain an insight into the progress of the approximation from term to term, the writer has computed tables of $f(x)$ and of the corresponding approximations for some simple values of $n$. The numerical computations are much facilitated by the use of published tables of $\Phi_0(z)$ and of its first six derivatives. A set of such tables to seven decimal places for values of $z$ up to 4 at intervals of 0.01 are published in N. R. Jorgensen,[*] and a similar set of five-place tables to eight derivatives for $z$ up to 5 are published by James W. Glover.[†]

In my numerical computations, the ordinates $y = f(x)$ have ordinarily been found at the junction points described in § 3, because it is but natural to expect that the approximations might not be so good at these points as elsewhere. Thus, the values of $f(x)$ and of the first, second, and third approximations, given by one, two and three terms of (46) respectively, have been computed at junction points for $n = 12$. It turns out that at the point $x = 0$, the second approximation is not so near the value of $f(x)$ as the first approximation and that at the three points $x = 4a, 5a,$ and $6a$, the third approximation is not so near the corresponding values of $f(x)$ as the second approximation. This simple case illustrates well the fact that the approximation is not improved at every point by the inclusion of an additional term, although it is improved in the sense of our least squares criterion.

---

[*] *Frequensflader og Korrelation*, 1916, pp. 177–193.
[†] *Tables for Applied Mathematics*, 1923, pp. 391-411.

UNIVERSITY OF IOWA,
   IOWA CITY, IOWA.