

The Definite Integral of the Product of Linear Functions

By R. E. von Holdt and R. J. Howerton

Part I. A common type of computation in applied science is evaluation of the definite integral of the product of two or more functions. Often, at least one of the functions in the product is phenomenological (i.e. pointwise determined from physical experiment) and it is not unusual to find that all the functions of the integrand product are such.

In the absence of proven mathematical models of the phenomena associated with the experimental determinations, one of several devices is commonly employed to obtain values of the function at unmeasured points. Least squares fitting to one of several classes of functions is one of the more common and is particularly useful (if weighted by the reciprocal of the experimental uncertainty) for purposes of gaining an insight to the nature of the physical phenomenon. When, however, the only immediate objective is to obtain functional values for integration purposes, a completely acceptable method is to approximate a smooth function by a minimum of line segments such that linear interpolation between successive end-points will not violate experimental error. Whether a weighted least squares fit or a linear approximation is used, it is obvious that at almost all points the physical function will not be represented correctly. Just how incorrect the resulting integrals or subsequent computations will be, due to errors in the functional representation of the integrands is the subject of many investigations. These investigations or sensitivity studies are extremely useful to establish requirements for maximum acceptable uncertainties in experimental determinations. For example, in certain measurements of phenomena associated with atomic nuclei, it is relatively easy to obtain values accurate to $\pm 10\%$ and almost impossible to improve accuracy beyond $\pm 2\%$. To obtain a 2% measurement requires many times the effort than to obtain a 10% measurement. Thus, one requesting measurements must have a sound basis for requesting high accuracy experiments.

In the evaluation of definite integrals of phenomenologically determined functions, there are two sources of error. The first is associated with the uncertainties of the integrand and the second with the numerical integration scheme. Errors associated with numerical integration are subject to error analysis but, for the scientist, it would be extremely useful to limit the errors to a single class, namely, those associated with the uncertainties in the integrand. It is the purpose here to describe an exact method of integration applicable whenever the integrand can be represented as the product of " n " linear functions.

There are two cases to be considered. The first is immediately tractable and the second requires an intermediate step. Case I is defined as an integrand which is the product of empirically defined functions only. Case II is an integrand in which one or more of the factors of the product is analytically defined. For Case II it is simple to replace the analytically defined function or functions by a practical minimum

Received December 12, 1962, revised March 19, 1963.

of line segments such that linear interpolation between successive points yields intermediate values which do not deviate from the given function by more than a specified fraction [1]. Case II is thus reduced in form to Case I.

The statement and proof of the theorem appropriate to the case of "n" linear functions is presented in Part II of this paper. For this discussion the example of the definite integral of the product of three linear functions is presented without proof since the general proof is in Part II. If

$$\begin{aligned}
 f(x) &= ax + b \\
 g(x) &= cx + d \\
 h(x) &= ex + g
 \end{aligned}$$

$$\begin{aligned}
 \int_{x_1}^{x_2} f(x)g(x)h(x) dx &= \frac{x_2 - x_1}{4} \left[f(x_1)g(x_1)h(x_1) \right. \\
 &\quad \left. + \frac{f(x_1)g(x_1)h(x_2) + f(x_2)g(x_2)h(x_1)}{3} \right. \\
 (1) \quad &\quad \left. + \frac{f(x_1)g(x_2)h(x_1) + f(x_2)g(x_1)h(x_2) + f(x_1)g(x_2)h(x_2) + f(x_2)g(x_1)h(x_1)}{3} \right. \\
 &\quad \left. \left. + f(x_2)g(x_2)h(x_2) \right].
 \end{aligned}$$

In the special case: $h(x) \equiv 1$ the result is,

$$(2) \quad \frac{x_2 - x_1}{3} \left[f(x_1)g(x_1) + \frac{f(x_1)g(x_2) + f(x_2)g(x_1)}{2} + f(x_2)g(x_2) \right];$$

and if $g(x) \equiv h(x) \equiv 1$

$$(3) \quad \frac{x_2 - x_1}{2} [f(x_1) + f(x_2)].$$

If in (1), (2), (3), and (22), the product of linear functions is evaluated at $n - 1$ equally spaced points (producing n equal intervals), and the Newton-Cotes formulas of order n are applied, the numerical results will be identical since both processes are exact. Furthermore, as would be expected, our formulas can be converted to appropriate Newton-Cotes formulas through trivial but tedious manipulation. For more than one linear function our formula needs fewer arithmetic operations for its evaluation than the corresponding Newton-Cotes formula.

Part II. The Definite Integral of The Product of Linear Functions of One Variable.

THEOREM. Let $\{S(m)\}_{m=0}^n$ be such that

$$(1) \quad \prod_{i=1}^n \{(a + r_i) + y(b + r_i)\} \equiv \sum_{m=0}^n S(m)y^m,$$

and

$$(2) \quad C(n, m) = n!/(m!(n - m)!),$$

then

$$(3) \quad I_n = \int_a^b \prod_{i=1}^n (t + r_i) dt = \frac{b-a}{n+1} \sum_{m=0}^n S(m)/C(n, m).$$

Proof. Let $\{p(k)\}_{k=0}^n$ be such that

$$(4) \quad \prod_{i=1}^n (t + r_i) = \sum_{k=0}^n p(k)t^{n-k}$$

then

$$(5) \quad \begin{aligned} I_n &= \sum_{k=0}^n p(k) \frac{b^{n+1-k} - a^{n+1-k}}{n+1-k} \\ &= \frac{b-a}{n+1} \sum_{k=0}^n p(k) \sum_{l=0}^{n-k} a^{n-k-l} b^l \cdot \frac{n+1}{n+1-k}. \end{aligned}$$

From (1), $S(m)$ is a polynomial: homogeneous of degree n in $(a, b, r_1, r_2, \dots, r_n)$; of degree at most $(n - m)$ in a ; of degree at most m in b ; symmetric in (r_1, r_2, \dots, r_n) ; and linear in each of (r_1, r_2, \dots, r_n) . Hence $S(m)$ can be expressed as a linear polynomial in the $p(k)$'s in the form:

$$(6) \quad S(m) = \sum_{l=0}^m \sum_{k=m-l}^{n-l} g(m, k, l) p(k) a^{n-l-k} b^l.$$

Let $P(i, j)$ be the set of all (x_1, x_2, \dots, x_i) , where j of the x 's are equal to a and the remaining $(i - j)$ of the x 's are equal to b . Then

$$(7) \quad S(m) = \sum_{P(n, m)} \prod_{i=1}^n (x_i + r_i)$$

where the number of summands in (7) is $C(n, m)$.

For fixed k and l , $d(m, k, l)$ is the coefficient of

$$(8) \quad \prod_{i=1}^k r_i a^{n-l-k} b^l$$

in $S(m)$. Since each summand of (7) has exactly m x 's equal to b , the sub-sum of (7) over those summands which when expanded contain the term (8) is

$$(9) \quad \sum_{P(k, m-l)} \prod_{i=1}^k (x_i + r_i) \sum_{P(n-k, l)} \prod_{i=1}^{n-k} (x_i + r_{i+k})$$

where the undefined expressions for $k = 0, n$ are replaced by one. But $g(m, k, l)$ is the number of terms in (9) or

$$g(m, k, l) = C(k, m - l)C(n - k, l).$$

Using this in (6), we have

$$(10) \quad S(m) = \sum_{l=0}^m \sum_{k=m-l}^{n-l} C(k, m - l)C(n - k, l) p(k) a^{n-l-k} b^l.$$

Suppose that

$$(11) \quad I_n = \frac{b-a}{n+1} \sum_{m=0}^n f(m) S(m)$$

then

$$(12) \quad I_n = \frac{b-a}{n+1} \sum_{m=0}^n \sum_{l=0}^m \sum_{k=m-l}^{n-l} d(k, l, m)$$

where

$$d(k, l, m) = f(m)C(k, m-l)C(n-k, l)p(k)a^{n-l-k}b^l$$

From Figure 1,

$$I_n = \frac{b-a}{n+1} \sum_{l=0}^n \sum_{m=l}^n \sum_{k=m-l}^{n-l} d(k, l, m).$$

From Figure 2,

$$I_n = \frac{b-a}{n+1} \sum_{l=0}^n \sum_{k=0}^{n-l} \sum_{m=l}^{k+l} d(k, l, m)$$

From Figure 3,

$$I_n = \frac{b-a}{n+1} \sum_{k=0}^n \sum_{l=0}^{n-k} \sum_{m=l}^{k+l} d(k, l, m).$$

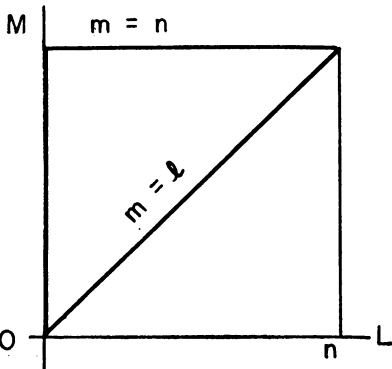


FIG. 1.

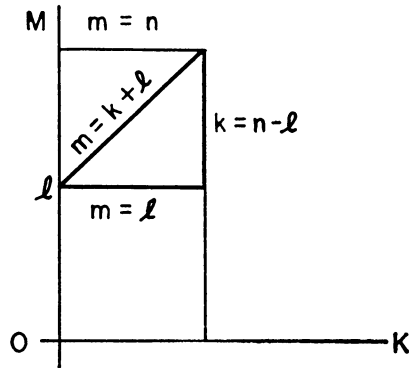


FIG. 2.

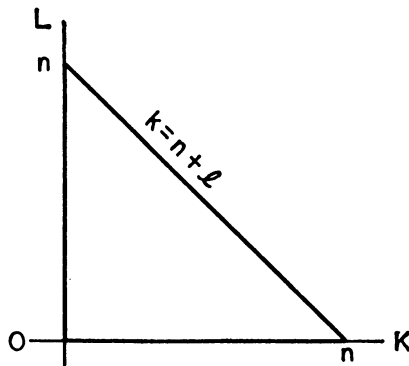


FIG. 3.

Thus (12) is

$$(13) \quad I_n = \frac{b-a}{n+1} \sum_{k=0}^n p(k) \sum_{l=0}^{n-k} a^{n-k-l} b^l C(n-k, l) \sum_{m=l}^{k+l} f(m) C(k, m-l)$$

Comparing (13) with (5), we see that (11) holds if and only if

$$(14) \quad \sum_{m=l}^{k+l} f(m) C(k, m-l) = \frac{n+1}{n+1-k} \frac{1}{C(n-k, l)},$$

$(l = 0, 1, \dots, n-k), (k = 0, 1, \dots, n).$

Setting $k = 0$ in (14), we have

$$(15) \quad f(l) = \frac{1}{C(n, l)}, \quad (l = 0, 1, \dots, n)$$

Before we show that (15) is a solution of (14) for $k \neq 0$, we note that

$$(16) \quad C(i, j) + C(i, j-1) = \frac{i!}{j!(i-j)!} + \frac{i!}{(j-1)!(i-j+1)!}$$

$$= \frac{i!}{j!(i-j+1)!} \cdot [(i-j+1) + j] = C(i+1, j)$$

and

$$(17) \quad \frac{1}{C(i, l)} + \frac{1}{C(i, l+1)} = \frac{l!(i-l)!}{i!} + \frac{(l+1)!(i-l-1)!}{i!}$$

$$= \frac{l!(i-l-1)!}{(i-1)!} \cdot \frac{[(i-l)+(l+1)]}{i} = \frac{i+1}{i} \frac{1}{C(i-1, l)}.$$

Suppose that (15) is a solution of (14) for $k = q$, or

$$(18) \quad \sum_{m=l}^{q+l} \frac{C(q, m-l)}{C(n, m)} = \frac{n+1}{n+1-q} \frac{1}{C(n-q, l)}, \quad (l = 0, 1, \dots, n-q).$$

Then for $(l = 0, 1, \dots, n-q-1)$,

$$\sum_{m=l}^{q+l+1} \frac{C(q+1, m-l)}{C(n, m)}$$

$$= \frac{1}{C(n, l)} + \sum_{m=l+1}^{q+l} \frac{C(q+1, m-l)}{C(n, m)} + \frac{1}{C(n, q+1+l)}$$

$$= \frac{1}{C(n, l)} + \sum_{m=l+1}^{q+l} \frac{C(q, m-l)}{C(n, m)} + \sum_{m=l+1}^{q+l} \frac{C(q, m-l-1)}{C(n, m)} + \frac{1}{C(n, q+1+l)}$$

$$= \sum_{m=l}^{q+l} \frac{C(q, m-l)}{C(n, m)} + \sum_{m=l+1}^{q+l+1} \frac{C(q, m-l+1)}{C(n, m)}$$

$$= \frac{n+1}{n+1-q} \left[\frac{1}{C(n-q, l)} + \frac{1}{C(n-q, l+1)} \right]$$

$$= \frac{n+1}{n+1-q} \cdot \frac{n-q+1}{n-q} \frac{1}{C(n-q-1, l)} = \frac{n+1}{n+1-q+1} \frac{1}{C(n-q+1, l)}$$

Thus if (15) is a solution of (14) for $k = q$, then (15) is a solution of (14) for $k = q + 1$. But (15) is the unique solution of (14) for $k = 0$, hence by mathematical induction (15) is the unique solution of (14). Hence (11) is equivalent to (15) and the theorem is proved.

Corollary.

$$(19) \quad \int_a^b \prod_{i=1}^n (c_i t + d_i) dt = \frac{(b-a)^{I[(n+1)/2]}}{n+1} \sum_{m=0}^{I[(n+1)/2]} \frac{S(m) + S(n-m)}{C(n, m)}$$

where $C(n, m) = n!/(m!(n-m)!)$, $I\left(\frac{n+1}{2}\right)$ is the integer part of $\left(\frac{n+1}{2}\right)$, and $S(m)$, ($m = 0, 1, \dots, n$), are defined by

$$(20) \quad \prod_{i=1}^n \{(c_i a + d_i) + y(c_i b + d_i)\} \equiv \sum_{m=0}^n S(m) y^m.$$

This follows upon setting $r_i = d_i/c_i$, $c_i \neq 0$ and multiplying by $\prod_{i=1}^n c_i$ in (1) and (3). For $c_i = 0$, we need only take the limit of (19) and (20) as $c_i \rightarrow 0$.

Let $a_i = c_i a + d_i$, $b_i = c_i b + d_i$, ($i = 1, 2, \dots, n$) and suppose that a_i and b_i are given instead of c_i and d_i , then $S(m)$ are defined by

$$(21) \quad \prod_{i=1}^n (a_i + y b_i) \equiv \sum_{m=0}^n S(m) y^m.$$

For $n = 1$, (19) reduces to the trapezoidal rule

$$\int_a^b (c_1 t + d_1) dt = \frac{b-a}{2} (a_1 + b_1).$$

For $n = 2$,

$$\begin{aligned} \int_a^b (c_1 t + d_1)(c_2 t + d_2) dt &= \frac{b-a}{3} \left[a_1 a_2 + \frac{a_1 b_2 + b_1 a_2}{2} + b_1 b_2 \right] \\ &= \frac{b-a}{6} [2a_1 a_2 + a_1 b_2 + b_1 a_2 + 2b_1 b_2] \\ &= \frac{b-a}{6} [a_1(a_1 + \overline{a_2 + b_2}) + b_1(\overline{a_2 + b_2} + b_2)] \\ &= \frac{b-a}{6} [a_1 a_1 + (a_1 + b_1)(a_2 + b_2) + b_1 b_2], \end{aligned}$$

and this is Simpson's Rule.

For $n > 2$, all of the $S(m)$ are obtained from (21) by only 2^{n+1} multiplications, and (19) may be rewritten as

$$(22) \quad \int_a^b \prod_{i=1}^n (c_i t + d_i) dt = \frac{b-a}{K(n)} \sum_{m=0}^{I[(n+1)/2]} e(n, m) [S(m) + S(n-m)]$$

where

n	$K(n)$	$e(n, 1)$	$e(n, 2)$	$e(n, 3)$	$e(n, 4)$	$e(n, 5)$
3	12	3	1			
4	60	12	3	2		
5	60	10	2	1		
6	420	60	10	4	3	
7	840	105	15	5	3	
8	2520	280	35	10	5	4
9	2520	252	28	7	3	2

Note that if $a_i \cdot b_i > 0 (i = 1, 2, \dots, n)$ then $S(m)$ is a sum of terms with all signs alike, thus the use of (21) with (19) involves no subtractions.

Note also that if some of the linear factors of the integrand are repeated the number of multiplications needed to evaluate all of the $S(m)$ may be reduced by applying the binomial theorem.

Lawrence Radiation Laboratory
 University of California
 Livermore, California

1. J. FORD & R. HOWERTON, *A FORTRAN Subroutine for Elimination of Superfluous Points of a Pointwise Determined Function*, URCL-7150, Lawrence Radiation Laboratory, 1962.