A METHOD OF DETERMINING THE CONSTANTS IN THE BIMODAL FOURTH DEGREE EXPONENTIAL FUNCTION

By

A. L. O'Toole

In a paper in this Journal[1] the present writer has discussed some of the mathematical properties of a class of definite integrals which arise in the study of the frequency function

$$(1) \qquad y = e^{-a^2(x^4+p_1x^3+p_2x^2+p_3x+p_4)}, \quad a \neq 0.$$

This function defines the system of frequency curves for which the method of moments is the best method of fitting[2]—i.e. best in the sense of maximum likelihood—and this fact gives importance to its study. The curves are typically bimodal, the nature and location of the modes being given by the roots of the equation

$$(2) \qquad 4x^3+3p_1x^2+2p_2x+p_3 = 0.$$

The first problem which arose was that of finding an expression for the value of the definite integral

$$(3) \qquad I_0 = \int_{-\infty}^{\infty} e^{-a^2(x^4+p_1x^3+p_2x^2+p_3x+p_4)}\,dx.$$

If x is replaced by $x - \frac{p_1}{4}$ this integral becomes

$$(4) \qquad I_0 = \int_{-\infty}^{\infty} e^{-a^2(x^4+px^2+qx+r)}\,dx,$$

[1]On the system of curves for which the method of moments is the best method of fitting. Vol. IV, No. 1, Feb. 1933, p. 1.

[2]R. A. Fisher, On the mathematical foundations of theoretical statistics, Philosophical Transactions of the Royal Society of London, vol. 222, series A (1921), p. 355.

or

$$(5)\quad I_0 = k\int_{-\infty}^{\infty} e^{-a^2(x^4+px^2+qx)}dx \quad \text{where} \quad k = e^{-a^2r},$$

or, replacing $x\sqrt{a}$ by x where a is the positive square root of a^2

$$(6)\quad I_0 = \frac{k}{\sqrt{a}}\int_{-\infty}^{\infty} e^{-(x^4+apx^2+a^{\frac{3}{2}}qx)}dx,$$

or

$$(7)\quad I_0 = K\int_{-\infty}^{\infty} e^{-(x^4+Px^2+Qx)}dx,$$

where $\quad K = \frac{k}{\sqrt{a}}, \quad P = ap, \quad Q = a^{\frac{3}{2}}q.$

No real loss of generality is incurred in studying (5), (6) or (7) rather than (3). For the purposes of the previous paper it was found convenient to discuss certain special cases of (7) first, then (7) itself and later (5). Having in mind the practical purposes of this note, however, attention will be focused first on the form (5) and afterwards on (3). The transformations from the expressions obtained in the previous paper are very simple. For (5) the special cases studied and a few of the more important results obtained may be stated here as follows:
Type I:

$p = q = 0.$

$$I_0 = k\int_{-\infty}^{\infty} e^{-a^2x^4}dx = \frac{k}{\sqrt{2a}}\Gamma\left(\frac{1}{4}\right),$$

$$I_{2n} = k\int_{-\infty}^{\infty} x^{2n}e^{-a^2x^4}dx = \frac{k}{2a^{(2n+1)/2}}\Gamma\left(\frac{2n+1}{4}\right),\ n = 0, 1, 2, 3, \ldots,$$

$$I_{2n+1} = k\int_{-\infty}^{\infty} x^{2n+1}e^{-a^2x^4}dx = 0, \quad n = 0, 1, 2, 3, \ldots,$$

$$u_1 = \frac{I_1}{I_0} = 0,$$

$$u_2 = \frac{I_2}{I_0} = \frac{\Gamma(\frac{3}{4})}{a\Gamma(\frac{1}{4})},$$

$$u_3 = \frac{I_3}{I_0} = 0,$$

$$u_4 = \frac{I_4}{I_0} = \frac{\Gamma(\frac{5}{4})}{a^2\Gamma(\frac{1}{4})} = \frac{1}{4a^2},$$

hence

$$a^2 = \frac{1}{4u_4},$$

$$u_{2n} = \frac{I_{2n}}{I_0} = \frac{\Gamma(\frac{2n+1}{4})}{a^n\Gamma(\frac{1}{4})}, \quad n = 0, 1, 2, 3, \ldots\ldots,$$

$$u_{2n+1} = \frac{I_{2n+1}}{I_0} = 0, \quad n = 0, 1, 2, 3, \ldots\ldots .$$

Obviously, of course, k depends upon the total frequency and hence if the total frequency is

$$k = \frac{N}{\int_{-\infty}^{\infty} e^{-a^2x^4}dx} = \frac{2N\sqrt{a}}{\Gamma(\frac{1}{4})}.$$

This curve has a single mode located at $x=0$ and is symmetrical with respect to the ordinate at $x=0$.

Type II:

$$q=0,\ p=-2b,\ b>0,$$

$$I_0=k\int_{-\infty}^{\infty}e^{-a^2(x^4-2bx^2)}dx$$

$$=\frac{k}{\sqrt{a}}\int_{-\infty}^{\infty}e^{-(x^4-2abx^2)}dx$$

$$=\frac{k}{\sqrt{a}}\Big[\frac{1}{2}\Gamma(\tfrac{1}{4})(1+\frac{a^2b^2}{2!}+\frac{5\cdot 1a^4b^4}{4!}+\frac{9\cdot 5\cdot 1a^6b^6}{6!}+\frac{13\cdot 9\cdot 5\cdot 1a^8b^8}{8!}+\cdots)$$
$$+ab\Gamma(\tfrac{3}{4})(1+\frac{3a^2b^2}{3!}+\frac{7\cdot 3\cdot a^4b^4}{5!}+\frac{11\cdot 7\cdot 3a^6b^6}{7!}+\cdots\cdots)\Big]$$

$$=\frac{k}{\sqrt{a}}\Big[e^{\frac{a^2b^2}{2}}\frac{1}{2}\Gamma(\tfrac{1}{4})(1+\frac{a^4b^4}{3\cdot 4}+\frac{a^8b^8}{7\cdot 3\cdot 4^2\cdot 2!}+\cdots\cdots)$$
$$+ab\Gamma(\tfrac{3}{4})(1+\frac{a^4b^4}{5\cdot 4}+\frac{a^8b^8}{9\cdot 5\cdot 4^2\cdot 2!}+\cdots\cdots)\Big]$$

It was shown that this integral could be expressed in terms of the Bessel functions $J_{\frac{1}{4}}$ and $J_{-\frac{1}{4}}$ as follows:

$$I_0=\frac{k}{\sqrt{a}}\left[(\frac{a^2b^2}{2})^{\frac{1}{4}}e^{\frac{a^2b^2}{2}}\left\{AJ_{\frac{1}{4}}(\frac{-ia^2b^2}{2})+BJ_{-\frac{1}{4}}(\frac{-ia^2b^2}{2})\right\}\right]$$

where

$$A=\frac{2^{\frac{3}{4}}\Gamma(\frac{1}{4})\Gamma(\frac{3}{4})}{\sqrt[4]{-i}},$$

$$B=\frac{\frac{1}{2}\Gamma(\frac{1}{4})\Gamma(\frac{3}{4})}{\sqrt[4]{2i}},\quad i=\sqrt{-1}.$$

If the total frequency is N

$$k = \frac{N}{\int_{-\infty}^{\infty} e^{-a^2(x^4 - 2bx^2)} dx}$$

This curve is symmetrical with respect to the ordinate at $x = 0$ and has two real modes located at $x = \pm\sqrt{b}$.

Type III: $p = 0$.

$$I_0 = k\int_{-\infty}^{\infty} e^{-a^2(x^4 + qx)} dx$$

$$= \frac{k}{\sqrt{a}}\int_{-\infty}^{\infty} e^{-(x^4 + a^{\frac{3}{2}}qx)} dx$$

$$= \frac{k}{2\sqrt{a}} \sum_{n=0}^{\infty} \frac{(a^{\frac{3}{2}}q)^{2n}}{(2n)!} \Gamma\left(\frac{2n+1}{4}\right)$$

$$= \frac{k}{2\sqrt{a}}\left[\Gamma\left(\tfrac{1}{4}\right)\left\{1 + \frac{(a^{\frac{3}{2}}q)^4}{4\cdot 4!} + \frac{5(a^{\frac{3}{2}}q)^8}{4^2\cdot 8!} + \frac{9\cdot 5(a^{\frac{3}{2}}q)^{12}}{4^3\cdot 12!} + \cdots\right\} + \Gamma\left(\tfrac{3}{4}\right)\left\{\frac{(a^{\frac{3}{2}}q)^2}{2!} + \frac{3(a^{\frac{3}{2}}q)^6}{4\cdot 6!} + \frac{7\cdot 3(a^{\frac{3}{2}}q)^{10}}{4^2\cdot 10!} + \cdots\right\}\right],$$

$$k = \frac{N}{\int_{-\infty}^{\infty} e^{-a^2(x^4 + qx)} dx}$$

This curve is not symmetrical and has only one real mode, that mode being located at x equal to the real cube root of negative q.

Type IV: The general case.

$$I_0 = k\int_{-\infty}^{\infty} e^{-a^2(x^4+px^2+qx)}dx$$

$$= \frac{k}{a}\int_{-\infty}^{\infty} e^{-(x^4+apx^2+a^{\frac{3}{2}}qx)}dx$$

$$= K\int_{-\infty}^{\infty} e^{-(x^4+Px^2+Qx)}dx.$$

It was shown that the value of this integral could be expressed as an infinite series each term of which involved two Bessel functions. But, as pointed out near the close of the previous paper, although this infinite series may be considered a theoretical solution of the problem, it does not lead to a simple method of determining the constants a^2, p, q, k which appear in the frequency function. It is the purpose of this note to give a practical method of determining these constants.

Beginning with (5)

$$I_0 = k\int_{-\infty}^{\infty} e^{-a^2(x^4+px^2+qx)}dx,$$

the *n-th* moment u'_n is defined by

$$u'_n = \frac{k\int_{-\infty}^{\infty} x^n e^{-a^2(x^4+px^2+qx)}dx}{k\int_{-\infty}^{\infty} e^{-a^2(x^4+px^2+qx)}dx} \tag{8}$$

$$= \frac{k}{I_0}\int_{-\infty}^{\infty} x^n e^{-a^2(x^4+px^2+qx)}dx.$$

Integrate I_0 by parts, letting $u = e^{-a^2(x^4+px^2)}$ and $dv = e^{-a^2qx}dx$. Then

$$(9) \qquad I_0 = -\frac{k}{q}\int_{-\infty}^{\infty}(4x^3+2px)e^{-a^2(x^4+px^2+qx)}dx.$$

Divide by I_0 and multiply by q and the result is

$$(10) \qquad q = -(4u_3' + 2pu_1').$$

Start again with I_0 in the form (5) and integrate by parts letting

$u = e^{-a^2(x^4+px^2+qx)}$ and $dv = dx$. Then

$$(11) \quad I_0 = ka^2\int_{-\infty}^{\infty}(4x^4+2px^2+qx)e^{-a^2(x^4+px^2+qx)}dx.$$

Divide by I_0 and then

$$1 = a^2(4u_4' + 2pu_2' + qu_1')$$

or

$$(12) \qquad a^2 = \frac{1}{4u_4' + 2pu_2' + qu_1'}.$$

Now integrate (11) by parts with $u = e^{-a^2(x^4+px^2+qx)}$ and

$dv = (4x^4+2px^2+qx)\,dx$. This leads to

$$(13) \qquad \begin{aligned} I_0 = \frac{ka^4}{30}\int_{-\infty}^{\infty}(96x^8 + 128px^6 + 84qx^5 + 40p^2x^4 \\ + 50pqx^3 + 15q^2x^2)e^{-a^2(x^4+px^2+qx)}dx. \end{aligned}$$

Divide by I_0 and obtain

$$1 = \frac{a^4}{30}(96u_8' + 128pu_6' + 84qu_5' + 40p^2u_4' + 50pqu_3' + 15q^2u_2')$$

or

$$(14)\quad a^4 = \frac{30}{96u'_8 + 128pu'_6 + 84qu'_5 + 40p^2u'_4 + 50pqu'_3 + 15q^2u'_2}.$$

Squaring in (12) the result is

$$(15)\quad a^4 = \frac{1}{(4u'_4 + 2pu'_2 + qu'_1)^2}$$

$$= \frac{1}{16u'^2_4 + 4p^2u'^2_2 + q^2u'^2_1 + 16pu'_2u'_4 + 8qu'_1u'_4 + 4pqu'_1u'_2}.$$

Eliminating a^4 between (14) and (15) the equation

$$(16)\quad p^2(40u'_4 - 120u'^2_2) + q^2(15u'_2 - 30u'^2_1) + pq(50u'_3 - 120u'_1u'_2)$$
$$+ p(128u'_6 - 480u'_2u'_4) + q(84u'_5 - 240u'_1u'_4) + (96u'_8 - 48u'^2_4) = 0.$$

Using relation (10) and

$$(17)\qquad q^2 = 16u'^2_3 + 16pu'_1u'_3 + 4p^2u'^2_1$$

eliminate q from (16) obtaining

$$(18)\qquad 5Ap^2 + 2Bp + 2C = 0$$

and hence

$$(19)\qquad p = \frac{-B \pm \sqrt{B^2 - 10AC}}{5A}$$

where

$$(20)\quad \begin{cases} A = 2u'_4 - 6u'_2 + 15u'^2_1u'_2 - 6u'^4_1 - 5u'_1u'_3, \\ B = 90u'_1u'_2u'_3 - 60u'^3_1u'_3 - 25u'^2_3 + 16u'_6 - 60u'_2u'_4 - 21u'_1u'_5 + 60u'^2_1u'_4, \\ C = 30u'_2u'^2_3 - 60u'^2_1u'^2_3 - 42u'_3u'_5 + 120u'_1u'_3u'_4 + 12u'_8 - 60u'^2_4. \end{cases}$$

In order to decide upon one of the two values of p furnished by (18) notice that, equating the first derivative of the frequency function to zero, the location of the two modes and the minimum point between them is determined by the roots of the equation

$$4x^3+2px+q=0. \tag{21}$$

The condition for three real distinct roots in this equation is

(22) $-8p^3>27q^2$, which requires $p<0$,

where q is found from (17). If $-8p^3=27q^2$ then one of the modes coincides with the minimum point. If $p=q=0$ then both modes coincide with the minimum point.

Extracting the square root in (17) gives two values of q differing only in sign. Now it is easy to show either by geometrical considerations or by examining the algebraic manipulations leading to (18) that p is independent of the sign of q. Changing q to $-q$ in (5) has the same effect as changing x to $-x$ or, that is, reversing the order of the distribution and curve. Also, changing x to $-x$ leaves the even moments unaltered but changes the sign of every odd moment. Hence if the value of the function at the modal position on the left is greater than the value of the function at the modal position on the right then q is greater than zero. And if the value of the function at the modal position on the left is less than the value of the function at the modal position on the right then q is less than zero. If $q=0$ the curve is symmetrical with respect to the ordinate at $x=0$. Hence p and q are determined by (19), (17) and (22), the sign of q being fixed by examination of the data of the problem or, if necessary, by trial. The value of a^2 is then found by taking the positive square root in (15). Of course (14) would give the same value for a^2.

Now that a^2, p and q are determined, there remains only k to be found. If the total frequency is N then

$$k\int_{-\infty}^{\infty} e^{-a^2(x^4+px^2+qx)}dx = N$$

and hence

$$k = \frac{N}{\int_{-\infty}^{\infty} e^{-a^2(x^4+px^2+qx)}dx} \tag{23}$$

where the numerical value of the integral in the denominator can be found by mechanical quadrature to any desired degree of approximation. For purposes of the quadrature involved here it will be found that the simple rectangle quadrature formula will give as good results as could be desired.[8] Having found k then the constant r is also known since

$$\left[\begin{array}{l} k = e^{-a^2 r}, \\ \\ r = \dfrac{\log_e k}{-a^2} = \dfrac{\log_{10} k}{-a^2 \log_{10} e} \end{array}\right. \tag{24}$$

The points of inflexion are located by equating the second derivative of the function to zero. The equation is

$$a^2(4x^3+2px+q)^2 - 2(6x^2+p) = 0. \tag{25}$$

If now x be replaced by $x+m$ then

$$I_0 = \int_{-\infty}^{\infty} e^{-a^2(x^4+px^2+qx+r)}dx \tag{5}$$

becomes

$$I_0 = \int_{-\infty}^{\infty} e^{-a^2(x^4+p_1x^3+p_2x^2+p_3x+p_4)}dx \tag{3}$$

[8]On the degree of Approximation of Certain Quadrature Formulas. Annals of Mathematical Statistics, vol. IV, No. 2, May 1933, p. 143 by A. L. O'Toole.

where

$$(26)\quad \begin{cases} p_1 = 4m \\ p_2 = 6m^2 + p, \\ p_3 = 4m^3 + 2mp + q, \\ p_4 = m^4 + m^2p + mq + r. \end{cases}$$

The data[4] in the first two columns of the table given here will provide the basis for an illustration of the method described above for determining the constants. The numbers in the first column are the classes into which the plants were divided. In the second column are found the frequencies corresponding to the various classes. In constructing the third column the origin for x was arbitrarily placed to correspond to the class 25. Taking

$$u'_n = \frac{\Sigma x^n f(x)}{\Sigma f(x)}$$

the first six moments and the eighth moment are found to be

$$u'_1 = \frac{88}{452} = 0.1946903,$$

$$u'_2 = \frac{14086}{452} = 31.16372,$$

$$u'_3 = \frac{12248}{452} = 27.09735,$$

$$u'_4 = \frac{1000264}{452} = 2212.973,$$

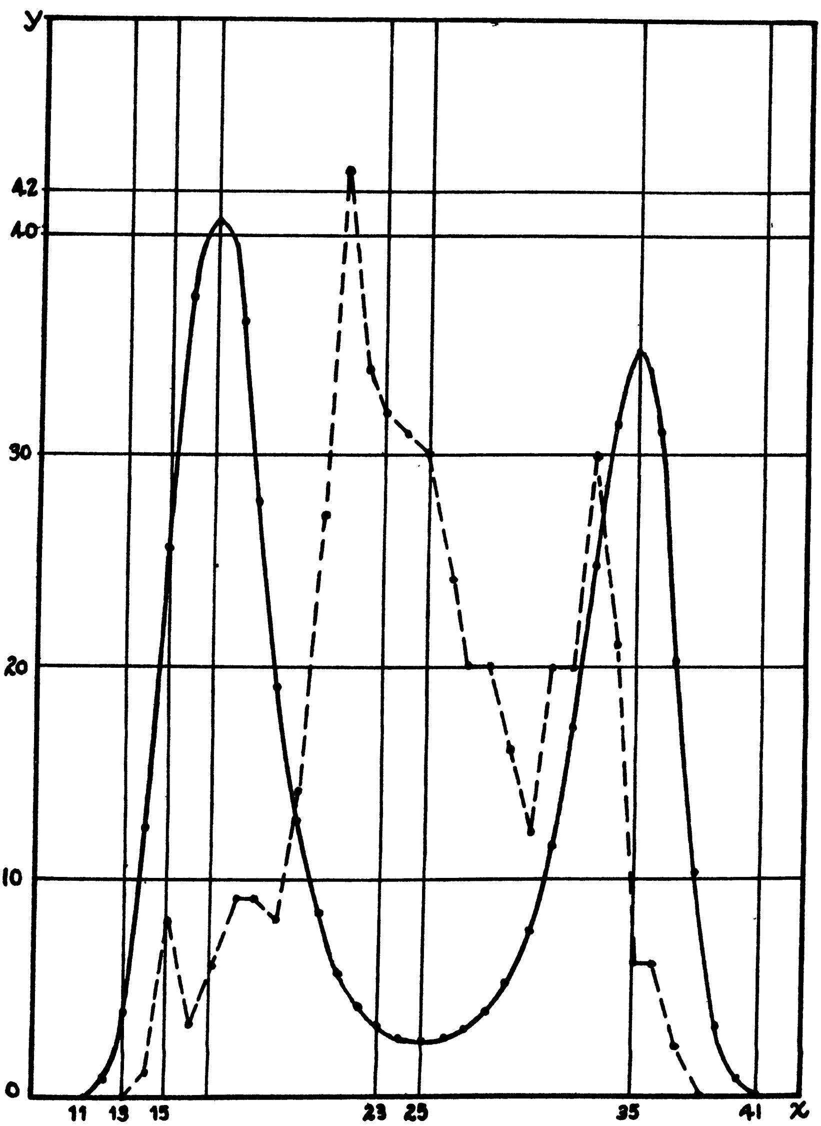
$$u'_5 = \frac{185480}{452} = 410.3540,$$

$$u'_6 = \frac{94571296}{452} = 209228.5,$$

$$u'_8 = \frac{10428472504}{452} = 23071842.$$

[4]This data, except for slight modifications, was extracted from that of W. L. Tower on the Seriation of Counts of Rays of Chrysanthemum Leucanthemum, Biometrika No. 1, 1901-2, p. 313.

CLASS	FREQUENCY $f(x)$	x	y'	$2.5y'=y$	COMPUTED FREQUENCY
9		-16	.031102	.077755	0
10		-15	.300472	.751180	1
11		-14	1.583594	3.958985	4
12	1	-13	4.991068	12.477670	12
13	8	-12	10.246676	25.616690	26
14	3	-11	14.831798	37.079495	37
15	6	-10	16.280112	40.700280	41
16	9	- 9	14.482540	36.206350	36
17	9	- 8	11.089036	27.722590	28
18	8	- 7	7.712195	19.280487	19
19	14	- 6	5.108903	12.772257	13
20	27	- 5	3.359072	8.397680	8
21	43	- 4	2.269766	5.674415	6
22	34	- 3	1.621764	4.054410	4
23	32	- 2	1.252760	3.131900	3
24	31	- 1	1.062910	2.657275	3
25	30	0	1.000000	2.500000	3
26	24	1	1.046530	2.616325	3
27	20	2	1.214445	3.036112	3
28	20	3	1.547935	3.869837	4
29	16	4	2.133051	5.332627	5
30	12	5	3.108096	7.770240	8
31	20	6	4.654336	11.635840	12
32	20	7	6.917724	17.294310	17
33	30	8	9.793409	24.483522	24
34	21	9	12.593310	31.483275	31
35	6	10	13.938151	34.845377	35
36	6	11	12.502565	31.256412	31
37	2	12	8.504388	21.260970	21
38		13	4.078577	10.196442	10
39		14	1.274131	3.185327	3
40		15	.238029	.595072	1
41		16	.024259	.060647	0
	452		180.792704		452


y
42
40
30
20
10
0
11
13
15
23
25
35
41
x

Formulas (20) give

$$A = -1409.786$$
$$B = -790428.9,$$
$$C = -15354106.$$

Hence from (19)

$$p = -202.7862$$

or

$$p = -21.48292.$$

But $p = -21.48292$ and the value of q to which it leads do not satisfy the relation (22) hence use $p = -202.7862$. Calculate q from (17) and use the positive square root since an examination of the data shows that the value of the function at the left modal value is greater than the value of the function at the right modal value. Hence

$$q = 29.4284.$$

Formula (15) now gives as the positive square root

$$a^2 = 0.0002640.$$

Using these values for a^2, p, q the values of the function

$$y' = e^{-a^2(x^4 + px^2 + qx)}$$

are calculated for integral values of x from $x = -16$ to $x = 16$ and tabulated in column four. The constant k is then found by dividing the total frequency 452 by the sum of column four. Hence

$$k = \frac{452}{180.792704} = 2.500100.$$

By (24)

$$r = -3472.578.$$

The function can now be written

$$y = 2.5e^{-.0002640(x^4 - 202.7862x^2 + 29.4284x)} \text{ taking } k = 2.5,$$

or $y = e^{-.0002640(x^4 - 202.7862x^2 + 29.4284x - 3472.578)}$.

The values of the ordinates for this function are given in column five and to the nearest integer in column six.

Equation (21) becomes

$$4x^3-2(202.7862)x+29.4284=0$$

which has the roots (approximate) $x=-10.1$, $x=0.07$, $x=10.03$ It should be noted that the sum of the three roots must equal zero. Hence the modes are located at $x=-10.1$ and at $x=10.03$ with the minimum point at $x=0.03$. These roots can be determined to any desired number of decimal places by Horner's method.

If now x is replaced by $x=-25$ so that the new values of x are respectively equal to the numbers in the class column, the function becomes

$$y=e^{-.0002640(x^4-100x^3+3547.214x^2-52331.26x+259605.3)}$$

The modes are now located at $x=14.9$ and $x=35.03$ with the minimum point at $x=25.07$. In the figure are shown the original distribution and the curve represented by this equation.

A. L. O'Toole