EQUIMODAL FREQUENCY DISTRIBUTIONS

By

Edwin D. Mouzon, Jr.

The object of this paper is the determination of a set of frequency curves, each of which will give a better fit to the modal neighborhood of the data to which it is applied than is often found in the existing methods. Interest in this subject was aroused in the following way. First it was discovered that a great number of distributions of data derived from a study of the financial ratios of public utility companies conformed to the same general type of curve. Second, it developed that the type of curve designated by the Pearsonian criterion quite often yielded a very poor fit to the data. The mode determined by the theoretical curve was obviously unsuited to the actual data. Furthermore, in some cases, on the left extremity of the distribution, the rise of the curve to the mode was too steep for a good fit. The accompanying chart (p. 140) presents a particular instance of these conditions, together with the curve fitted by the method developed in this paper.

The curves which were used in this study of financial ratios were those developed by Pearson and Elderton from a consideration of the various cases which arose in the solution of the differential equation

$$\frac{dy}{dx} = \frac{y(x-a)}{F(x)}$$

where $F(x)$ was assumed to be expansible in ascending powers of x. The other assumptions made were that $F(x) = b_0 + b_1 x + b_2 x^2$, and that the constants a, b_0, b_1, and b_2 were determined by equating the moments of the raw data to the moments of the theoretical distribution. Here we will modify these assumptions, and under the new conditions determine the principal types of curves which arise when the polynomial in the denominator is of the third or lower degree.

The new assumption is that the value of the constant, a, the mode, is determined first from the observed data, and equated to the value of the mode in the theoretical distribution. This method of procedure is particularly adapted to economic data, as it assures a good fit about the mode, notwithstanding the fact that in some raw data the mode is a rather vague concept. The fit about the mode is of primary importance in much economic data.

II. We begin with the case of the cubic in the denominator, that is with the differential equation

$$\frac{dy}{dx} = \frac{(x-a)\,y}{b_0 + b_1 x + b_2 x^2 + b_3 x^3}\text{ , or}$$

$$(b_0 + b_1 x + b_2 x^2 + b_3 x^3)\,\frac{dy}{dx} = y\,(x-a)$$

where a is known. Multiplying both sides by x^n, integrating, and using the notation $\mu'_n = \int y\,x^n\,dx$, we have

$$n b_0 \mu'_{n-1} + (n+1)\, b_1 \mu'_n + (n+2)\, b_2 \mu'_{n+1} + (n+3)\, b_3 \mu'_{n+2} = a\mu'_n - \mu'_{n+1}$$

Putting $n = 0, 1, 2, 3$, and changing the origin to the mean, we have

$$(1)\qquad \begin{cases} 0 b_0 + b_1 + 0 b_2 + 3\mu_2 b_3 = a \\ b_0 + 0 b_1 + 3\mu_2 b_2 + 4\mu_3 b_3 = -\mu_2 \\ 0 b_0 + 3\mu_2 b_1 + 4\mu_3 b_2 + 5\mu_4 b_3 = a\mu_2 - \mu_3 \\ 3\mu_2 b_0 + 4\mu_3 b_1 + 5\mu_4 b_2 + 6\mu_5 b_3 = a\mu_3 - \mu_4 \end{cases}$$

Solving these equations for b_0, b_1, b_2, and b_3, we have

$$b_0 = \frac{\begin{vmatrix} a & 1 & 0 & 3\mu_2 \\ -\mu_2 & 0 & 3\mu_2 & 4\mu_3 \\ a\mu_2 - \mu_3 & 3\mu_2 & 4\mu_3 & 5\mu_4 \\ a\mu_3 - \mu_4 & 4\mu_3 & 5\mu_4 & 6\mu_5 \end{vmatrix}}{\begin{vmatrix} 0 & 1 & 0 & 3\mu_2 \\ 1 & 0 & 3\mu_2 & 4\mu_3 \\ 0 & 3\mu_2 & 4\mu_3 & 5\mu_4 \\ 3\mu_2 & 4\mu_3 & 5\mu_4 & 6\mu_5 \end{vmatrix}} = \frac{A}{\Delta}$$

$$b_1 = \frac{\begin{vmatrix} 0 & a & 0 & 3\mu_2 \\ 1 & -\mu_2 & 3\mu_2 & 4\mu_3 \\ 0 & a\mu_2-\mu_3 & 4\mu_3 & 5\mu_4 \\ 3\mu_2 & a\mu_3-\mu_4 & 5\mu_4 & 6\mu_5 \end{vmatrix}}{\Delta} = \frac{B}{\Delta}$$

$$b_2 = \frac{\begin{vmatrix} 0 & 1 & a & 3\mu_2 \\ 1 & 0 & -\mu_2 & 4\mu_3 \\ 0 & 3\mu_2 & a\mu_2-\mu_3 & 5\mu_4 \\ 3\mu_2 & 4\mu_3 & a\mu_3-\mu_4 & 6\mu_5 \end{vmatrix}}{\Delta} = \frac{C}{\Delta}$$

$$b_3 = \frac{\begin{vmatrix} 0 & 1 & 0 & a \\ 1 & 0 & 3\mu_2 & -\mu_2 \\ 0 & 3\mu_2 & 4\mu_3 & a\mu_2-\mu_3 \\ 3\mu_2 & 4\mu_3 & 5\mu_4 & a\mu_3-\mu_4 \end{vmatrix}}{\Delta} = \frac{D}{\Delta}$$

The differential equation then becomes

$$\frac{dy}{y} = \frac{(x-a)\,dx}{\frac{A}{\Delta} + \frac{Bx}{\Delta} + \frac{Cx^2}{\Delta} + \frac{Dx^3}{\Delta}}$$

The solution of the differential equation depends on the nature of the zeros of the denominator of the right hand member, that is on the discriminant of the general cubic,

$$18\, b_3 b_2 b_1 b_0 - 4 b_2^3 b_0 + b_2^2 b_1^2 - 4 b_2 b_1^3 - 27 b_3^2 b_0^2$$

The cubic has three distinct real zeros, one real and two imaginary zeros, or at least two real and equal zeros, according as the discriminant is greater than zero, less than zero, or equal to zero. We will expect, therefore, three general types of curves when the integration is effected.

III. If we assume that $b_3 = 0$, we will have only three constants, b_0, b_1, and b_2 to determine, and the equations (1) become

$$
(2) \quad \begin{cases} b_1 = a \\ b_0 \quad \cdot \quad + 3b_2\mu_2 = -\mu_2 \\ 3b_1\mu_2 + 4b_2\mu_3 = -\mu_3 + a\mu_2 \end{cases}
$$

Solving these equations simultaneously, we find

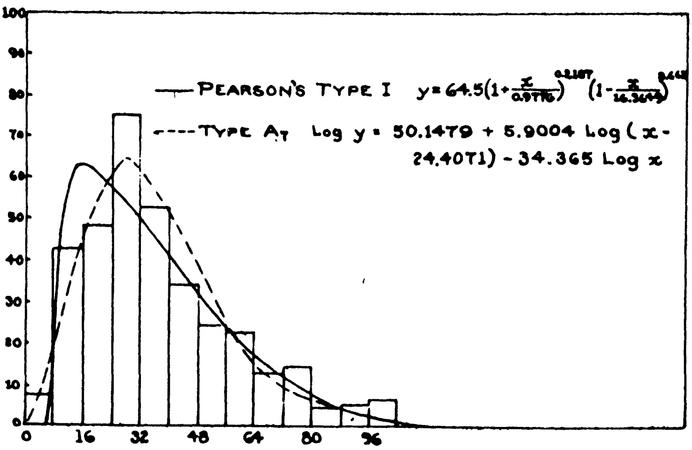
$$
b_0 = \frac{-\mu_2\mu_3 + 6a\mu_2^2}{4\mu_3}
$$

$$
b_1 = a
$$

$$
b_2 = \frac{-\mu_3 - 2a\mu_2}{4\mu_3}
$$

Thus, in the case of the quadratic in the denominator, we have determined the constants in terms of the mode, and the first, second, and third moments of the raw data. In other words, we are calculating the theoretical curve under the assumption that its mode, mean, standard deviation, and skewness are equal respectively to the mode,

Ratio of Revenue to Net Worth in 351 Traction Companies



mean, standard deviation, and skewness of the raw data. The differential equation then becomes

$$\frac{1}{y}\frac{dy}{dx}=\frac{x-a}{\dfrac{-\mu_2\mu_3+6a\mu_2^2}{4\mu_3}+ax+\dfrac{-\mu_3-2a\mu_2}{4\mu_3}x^2}$$

Now, the solution of this differential equation depends on the particular values of the constants in the denominator, i. e., on the quadratic discriminant $b_1^2-4b_0b_2$ Again, we will expect three general types of curves when the integration is effected.

If we assume $b_3=b_2=0$, equations (1) become

$$b_1=a$$

(3)

$$b_0=-\mu_2$$

and the differential equation is

$$\frac{1}{y}\frac{dy}{dx}=-\frac{x-a}{\mu_2-ax}$$

If we keep only b_0, we have $b_0=-\mu_2$ and $a=0$, and our equation is $\frac{1}{y}\frac{dy}{dx}=-\frac{x}{\mu_2}$

We now turn to a discussion of the various types of curves which arise from the solution of the preceding differential equations. The following classification will be made—Class A will include all curves arising from the solution of differential equations in which $F(x)$ has real and unequal zeros. Class B will include all curves arising from the solution of differential equations in which $F(x)$ has complex zeros. Class C will include all curves arising from the solution of differential equations in which $F(x)$ has at least two equal zeros.

Type A-1

IV. When all the zeros are positive, the differential equation may be written in the form

$$\frac{dy}{y}=\frac{(x-a)dx}{b_3(x-A_1)(x-A_2)(x-A_3)}$$

where we assume $A_1 > A_2 > A_3$. Separating into partial fractions and integrating, we have

$$\begin{aligned}\log y = &\frac{(A_2-A_3)(A_1-a)}{b_3(A_1-A_3)(A_2-A_3)(A_1-A_2)}\log(x-A_1)\\ &-\frac{(A_1-A_3)(A_2-a)}{b_3(A_1-A_3)(A_2-A_3)(A_1-A_2)}\log(x-A_2)\\ &+\frac{(A_1-A_2)(A_3-a)}{b_3(A_1-A_3)(A_2-A_3)(A_1-A_2)}\log(x-A_3)\\ &+\log k\end{aligned}$$

Exponentiating

$$y=\frac{k(x-A_1)^{\frac{(A_2-A_3)(A_1-a)}{b_3(A_1-A_3)(A_2-A_3)(A_1-A_2)}}(x-A_3)^{\frac{(A_1-A_2)(A_3-a)}{b_3(A_1-A_3)(A_2-A_3)(A_1-A_2)}}}{(x-A_2)^{\frac{(A_1-A_3)(A_2-a)}{b_3(A_1-A_3)(A_2-A_3)(A_1-A_2)}}}$$

Transferring the origin to the mode, i. e., putting x for $x-a$ we have

$$y=\frac{k(x+a-A_1)^{\frac{c_1a_1}{s}}(x+a-A_3)^{\frac{c_3a_3}{s}}}{(x+a-A_2)^{\frac{c_2a_2}{s}}}$$

$$=\frac{y_0(1-x/a_1)^{\frac{c_1a_1}{s}}(1-x/a_3)^{\frac{c_3a_3}{s}}}{(1-x/a_2)^{\frac{c_2a_2}{s}}}$$

where

$$c_1=A_2-A_3 \qquad c_2=A_1-A_3 \qquad c_3=A_1-A_2$$

$$a_1=A_1-a \qquad a_2=A_2-a \qquad a_3=A_3-a$$

$$s=b_3(A_1-A_2)(A_1-A_3)(A_2-A_3)$$

Then $c_2a_2=c_1a_1+c_3a_3$

Let $\frac{c_1}{s}=m_1$, $\frac{c_2}{s}=m_2$, $\frac{c_3}{s}=m_3$

Then $m_2=m_1-m_3$

The equation now becomes

$$y = \frac{y_0 (1 - x/a_1)^{m_1 a_1} (1 - x/a_3)^{m_3 a_3}}{(1 - \frac{x}{a_2})^{m_2 a_2}}$$

With the exception of y_0, the values of all the constants in the equation are known in terms of moments. Two methods will be given for its determination.

First—Calculate the area under the curve, using the theoretical ordinates measured in terms of y_0. Let this area equal N, the number of observations, and solve for y_0.

Thus $y = \frac{N}{B_1 + B_2 + \cdots + B_n}$, where $y_0 B_i$ represents the areas calculated from the theoretical ordinates.

Second—Calculate the value of χ^2 (Chi-square) with the theoretical ordinates measured in terms of y_0. Set the first derivative of this expression equal to zero, and determine the value of y_0 which makes χ^2 a minimum. From the goodness of fit point of view, this gives the best possible value of y_0.

Thus $\chi^2 = \sum \frac{(y_0 B_i - O_i)^2}{y_0 B_i}$, where the $y_0 B_i$ represent the theoretical areas as before, and the O_i represent the observed areas.

Setting the first derivative of this expression equal to zero, we have

$$\sum B_i - \frac{1}{y_0^2} \sum \frac{O_i^2}{B_i} = 0$$

$$y_0^2 = \frac{\sum \frac{O_i^2}{B_i}}{\sum B_i}$$

Type A-2

V. When there are two positive zeros and one negative zero, the equation may be written in the form

$$\frac{dy}{y} = \frac{(x-a)\,dx}{b_3\,(x-A_1)(x-A_2)\,(x+A_3)}, \quad \text{where } A_1 > A_2 > -A_3$$

Proceeding as in Type A-I, we find

$$y = \frac{y_0\left(1-x/a_1\right)^{\frac{c_1 a_1}{s}}}{\left(1-x/a_2\right)^{\frac{c_2 a_2}{s}}\left(1-x/a_3\right)^{\frac{c_3 a_3}{s}}}$$

where

$$c_1 = A_2 + A_3 \qquad a_1 = A_1 - a$$
$$c_2 = A_1 + A_3 \qquad a_2 = A_2 - a \qquad s = b_3 c_1 c_2 c_3$$
$$c_3 = A_1 - A_2 \qquad a_3 = A_3 + a$$

y_0 is calculated as in Type A-I. The origin is at the mode.

Type A-3

VI. When there are two negative zeros and one positive zero, the equation may be written in the form

$$\frac{dy}{y} = \frac{(x-a)\,dx}{b_3\,(x-A_1)(x+A_2)(x+A_3)}, \quad \text{where } A_1 > -A_2 > -A_3$$

Using the same method as before, we find

$$y = \frac{y_0\left(1-x/a_1\right)^{\frac{c_1 a_1}{s}}\left(1+x/a_2\right)^{\frac{c_2 a_2}{s}}}{\left(1+x/a_3\right)^{\frac{c_3 a_3}{s}}}, \quad \text{where}$$

$$a_1 = A_1 - a \qquad c_1 = -A_2 + A_3$$
$$a_2 = A_2 + a \qquad c_2 = A_1 + A_3 \qquad s = b_3 c_1 c_2 c_3$$
$$a_3 = A_3 + a \qquad c_3 = A_1 + A_2$$

y_0 is calculated as in the previous cases. The origin is at the mode.

Type A-4

VII. Where all three zeros are negative, the equation may be written in the form

$$\frac{dy}{y} = \frac{(x-a)\,dx}{b_3(x+A_1)(x+A_2)(x+A_3)}, \text{ where } -A_1 > -A_2 > -A_3$$

Proceeding as in the last three cases, we find

$$y = \frac{y_0\left(1+x/a_2\right)^{\frac{c_2 a_2}{s}}}{\left(1+x/a_1\right)^{\frac{c_1 a_1}{s}}\left(1+x/a_3\right)^{\frac{c_3 a_3}{s}}}, \text{ where}$$

$$\begin{array}{lll} a_1 = A_1 + a & c_1 = -A_2 + A_3 & \\ a_2 = A_2 + a & c_2 = -A_1 + A_3 & s = b_3 c_1 c_2 c_3 \\ a_3 = A_3 + a & c_3 = -A_1 + A_2 & \end{array}$$

y_0 is determined as in the previous cases. The origin is at the mode.

Type A-5

VIII. In Type A-3, suppose $A_1 = A_2$. Then

$$y = \frac{y_0\left(1 - x/(A_1 - a)\right)^{\frac{(-A_1+A_3)(A_1-a)}{s}}\left(1 + x/(A_1+a)\right)^{\frac{(A_1+A_3)(A_1+a)}{s}}}{\left(1 + \frac{x}{A_3+a}\right)^{\frac{2A_1(A_3+a)}{s}}}$$

where $s = b_3\ (A_1+A_3)(-A_1+A_3)\ (2A_1)$

Suppose a, the mode, is at the mean, that is, is equal to zero. Then

$$A_1 - a = A_1 + a = a_1$$

Then

$$y = \frac{y_0 \left(1 - x/a_1\right)^{\frac{(-A_1+A_3)\,a_1}{s}} \left(1 + x/a_1\right)^{\frac{(A_1+A_3)\,a_1}{s}}}{\left(1 + x/a_3\right)^{\frac{2A_1 a_3}{s}}}$$

y_0 is calculated as before. The origin is at the mode.

Type A-6

IX. In Type A-3, suppose one of the zeros is zero, say A_1. The equation then becomes

$$y = \frac{k x^{\frac{(-A_2+A_3)(-a)}{b_3 A_2 A_3 (-A_2+A_3)}} \left(x+A_2\right)^{\frac{A_3(A_2+a)}{b_3 A_2 A_3 (-A_2+A_3)}}}{\left(x+A_3\right)^{\frac{A_2(A_3+a)}{b_3 A_2 A_3 (-A_2+A_3)}}}$$

$$= \frac{y_0 x^{\frac{-a}{b_3 A_2 A_3}} \left(1 + x/A_2\right)^{\frac{A_2+a}{b_3 A_2 (-A_2+A_3)}}}{\left(1 + x/A_3\right)^{\frac{A_3+a}{b_3 A_3 (-A_2+A_3)}}}$$

The values of all the constants except y_0 are known, and it may be calculated by either of the formulas given in Type A-1.

The origin is at the mean.

Type A-7

X. In case $F(x)$ is quadratic, and its zeros are of like sign, we have

$$\frac{dy}{y} = \frac{(x-a)\,dx}{b_2\,(x+A_1)(x+A_2)} = \frac{A_1+a}{b_2\,(A_1-A_2)}\,\frac{dx}{x+A_1} - \frac{A_2+a}{b_2\,(A_1-A_2)}\,\frac{dx}{x+A_2}$$

Integrating, we have

$$\log y = \frac{A_1+a}{b_2\,(A_1-A_2)} \log(x+A_1) + \frac{-(A_2+a)}{b_2\,(A_1-A_2)} \log(x+A_2) + \log y'$$

Exponentiating,

$$y = y'(x+A_1)^{\frac{A_1+a}{b_2(A_1-A_2)}}(x+A_2)^{\frac{-(A_2+a)}{b_2(A_1-A_2)}}$$

Changing x to $x - A_2$, we have

$$y = y'(x+A_1-A_2)^{\frac{A_1+a}{b_2(A_1-A_2)}}\, x^{\frac{-(A_2+a)}{b_2(A_1-A_2)}}$$

Let

$$A_1 - A_2 = -m \;;\quad \frac{A_1+a}{-b_2 m} = p_1 \;;\quad \frac{A_2+a}{b_2 m} = p_2$$

Then

$$y = y_0\,(x-m)^{p_1} x^{-p_2}$$

The constants A_1 and A_2 are given by the zeros of the quadratic in the denominator, and m, p_1, and p_2 are given in terms of these above. By integration of this equation between the limits m and ∞, it has been found[1] that

$$y_0 = \frac{N\Gamma(p_2)\, m^{p_2-p_1-1}}{\Gamma(p_1+1)\,\Gamma(p_2-p_1-1)}$$

y_0 may also be determined by finite integration by either of the methods given in Type A-1.

Origin = Mean $- A_2$

Type A-8

XI. This type occurs when $F(x)$ is a quadratic and the zeros are real and opposite in sign. The equation then becomes

$$\frac{dy}{y} = \frac{1}{b_2}\,\frac{x-a}{(x+A_1)(x-A_2)}$$

1. Elderton, "Frequency Curves and Correlation," p. 85.

Integrating, we obtain

$$\log y = \frac{1}{b_2}\,\frac{A_1+a}{A_1+A_2}\log(x+A_1) + \frac{1}{b_2}\,\frac{A_2-a}{A_1+A_2}\log(x-A_2) - \log y'$$

Exponentiating,

$$y = y'(x+A_1)^{\frac{1}{b_2}\frac{A_1+a}{A_1+A_2}}(x-A_2)^{\frac{1}{b_2}\frac{A_2-a}{A_1+A_2}}$$

Now, changing the origin to the mode, i. e., putting x for $x-a$, we have

$$y = y'(x+A_1+a)^{\frac{1}{b_2}\frac{A_1+a}{A_1+A_2}}(x-A_2+a)^{\frac{1}{b_2}\frac{A_2-a}{A_1+A_2}}$$

$$= y_0\left(1+\frac{x}{a_1}\right)^{m_1}\left(1-\frac{x}{a_2}\right)^{m_2}$$

where

$$a_1 = A_1 + a \qquad ; \qquad a_2 = A_2 - a$$

$$m_1 = \frac{1}{b_2}\,\frac{a_1}{A_1+A_2} \qquad ; \qquad m_2 = \frac{1}{b_2}\,\frac{a_2}{A_1+A_2}$$

and

$$\frac{m_1}{a_1} = \frac{m_2}{a_2}$$

The value of y_0 has been found[1] by integration to be

$$y_0 = \frac{N m_1^{m_1} m_2^{m_2}\,\Gamma(m_1+m_2+2)}{b(m_1+m_2)^{m_1+m_2}\,\Gamma(m_1+1)\,\Gamma(m_2+1)}$$

Where $b = a_1 + a_2$, N = total frequency. y_0 may also be calculated by either of the methods given in Type A-1. The origin is at the mode.

1. Elderton, "Frequency Curves and Correlation," p. 59.

TYPE A-9

XII. If the zeros of the quadratic are A_1, and $-A_1$, the equation may be written

$$\frac{dy}{y}=\frac{1}{b_2}\,\frac{A_1-a}{2A_1}\,\frac{dx}{x-A_1}+\frac{1}{b_2}\,\frac{A_1+a}{2A_1}\,\frac{dx}{x+A_1}$$

Integrating, we obtain

$$\log y=\frac{1}{b_2}\,\frac{A_1-a}{2A_1}\log(x-A_1)+\frac{1}{b_2}\,\frac{A_1+a}{2A_1}\log(x+A_1)+\log y$$

Exponentiating,

$$y=y'(x-A_1)^{\frac{1}{b_2}\frac{A_1-a}{2A_1}}(x+A_1)^{\frac{1}{b_2}\frac{A_1+a}{2A_1}}$$

Changing the origin to the mode, i. e., putting x for $x-a$, we have

$$y=y'(x+a-A_1)^{\frac{1}{b_2}\frac{A_1-a}{2A_1}}(x+a+A_1)^{\frac{1}{b_2}\frac{A_1+a}{2A_1}}$$

$$=y_0\left(1-\frac{x}{a_1}\right)^{m_1}\left(1+\frac{x}{c_1}\right)^{m_2}$$

where

$$a_1=A_1-a \qquad a_1+2a=A_1+a=c_1$$

$$\frac{a_1}{2b_2A_1}=m_1 \qquad \frac{c_1}{2b_2A_1}=m_2$$

Then, as before,

$$y_0=\frac{N}{a_1+c_1}\cdot\frac{m_1^{\,m_1}m_2^{\,m_2}\,\Gamma(m_1+m_2+2)}{(m_1+m_2)^{m_1+m_2}\,\Gamma(m_1+1)\,\Gamma(m_2+1)}$$

TYPE A-10

XIII. If $F(x)$ is assumed to be linear, the equation may be written

$$\frac{dy}{y} = \frac{x-a}{b_0+b_1x}\,dx = \left(\frac{1}{b_1} + \frac{-a-\frac{b_0}{b_1}}{b_1x+b_0}\right)dx$$

Integrating,

$$\log y = \frac{x}{b_1} + \frac{1}{b_1}\left(-a-\frac{b_0}{b_1}\right)\log(b_1x+b_0) + \log y'$$

Exponentiating,

$$y = y'\,e^{\frac{x}{b_1}}(b_1x+b_0)^{-\frac{1}{b_1}\left(a+\frac{b_0}{b_1}\right)}$$

Changing the origin to the mode by putting x for $x-a$, we have

$$y = y''e^{\frac{x+a}{b_1}}\left(1+\frac{x}{\frac{b_1a+b_0}{b_1}}\right)^{-\frac{1}{b_1}\left(\frac{b_1a+b_0}{b_1}\right)}$$

Now let

$$y''e^{\frac{a}{b_1}} = y_0 \quad ; \quad \frac{ab_1+b_0}{b_1} = m \quad ; \quad -\frac{1}{b_1} = \gamma$$

Then

$$y = y_0\,e^{-\gamma x}\left(1+\frac{x}{m}\right)^{\gamma m}.$$

The constants may be determined as follows:

When $b_2 = b_3 = 0$, it has been found that

$$b_0 = -\mu_2 \quad ; \quad b_1 = a$$

$$\therefore\ m = \frac{ab_1+b_0}{b_1} = \frac{a^2-\mu_2}{a}$$

$$\gamma = -\frac{1}{b_1} = -\frac{1}{a}$$

The value of y_0 has been found[1] to be

$$y_0 = \frac{Nq^{q+1}}{me^q \Gamma(q+1)}$$

where $q = \gamma m$.

y_0 may also be found by the methods of Type A-1. The origin is at the mode.

TYPE A-11 (The normal curve)

XIV. Putting $b_1 = b_2 = b_3 = 0$, we have

$$\frac{dy}{y} = \frac{x-a}{b_0}\,dx$$

Integrating,

$$\log y = \frac{x^2}{2b_0} - \frac{ax}{b_0} + \log c = \frac{(x-a)^2}{2b_0} + \log y'$$

Exponentiating,

$$y = y'e^{\frac{(x-a)^2}{2b_0}}$$

Changing the origin to the mode, and substituting the value for b_0 when $b_1 = b_2 = b_3 = 0$, that is $b_0 = -\mu_2$, we obtain

$$y = y_0\, e^{-\frac{x^2}{2\mu_2}}$$

To find the value of y_0 , integrate between the limits $-\infty$ and ∞ and find the total frequency N. It has been found[2] that

$$y_0 = \frac{N}{\sqrt{2\pi\mu_2}}$$

1. Elderton, "Frequency Curves and Correlation," p. 68.
2. Elderton, "Frequency Curves and Correlation," p. 91.

y_0 may also be found by either of the other two methods. The origin is at the mode.

TYPE B-1

XV. When $F(x)$ is a cubic, and two of the zeros are complex, the differential equation may be written in the form

$$\frac{dy}{y}=\frac{(x-a)\,dx}{b_3\,(x-A_1)\left[x^2-(A_2+A_3)\,x+A_2A_3\right]},\ \text{where } (A_2-A_3)^2<0$$

Separating into partial fractions and integrating,

$$\int\frac{dy}{y}=\frac{A_1-a}{b_3\left[A_1(A_1-A_2-A_3)+A_2A_3\right]}\int\frac{dx}{x-A_1}$$

$$-\frac{A_1-a}{b_3\left[A_1(A_1-A_2-A_3)+A_2A_3\right]}\int\frac{x\,dx}{x^2-(A_2+A_3)\,x+A_2A_3}$$

$$+\frac{a(A_1-A_2-A_3)+A_2A_3}{b_3\left[A_1(A_1-A_2-A_3)+A_2A_3\right]}\int\frac{dx}{x^2-(A_2+A_3)\,x+A_2A_3}$$

Now let $x'=x-\dfrac{A_2+A_3}{2}$; $M^2=\dfrac{-(A_2-A_3)^2}{2}$

$$k=a\,(A_1-A_2-A_3)+A_2A_3-\frac{(A_1-a)(A_2+A_3)}{2}$$

$$d=\left[A_1(A_1-A_2-A_3)+A_2A_3\right]b_3$$

Performing the integration, we have

$$\log y=\frac{A_1-a}{d}\log\left(x'+\frac{A_2+A_3}{2}-A_1\right)-\frac{A_1-a}{2d}\log\left(x'^2+M^2\right)$$
$$+\frac{k}{Md}\tan^{-1}\frac{x'}{M}+\log y_0$$

Exponentiating,

$$y=\frac{y_0\left[x'+\frac{1}{2}(A_2+A_3)-A_1\right]^{\frac{A_1-a}{d}}\,e^{\frac{k}{Md}\tan^{-1}\frac{x'}{M}}}{(x'^2+M^2)^{\frac{A_1-a}{2d}}}$$

$$=\frac{y_0\,(x+c)^{2m}\,e^{\frac{k}{Md}\tan^{-1}\frac{x'}{M}}}{(x'^2+M^2)^m}$$

where $m = \frac{A_1 - a}{2d}$, $\quad \frac{A_2 + A_3}{2} - A_1 = c$

y_0 may be determined as in Type A-1. Origin = Mean + $\frac{A_2+A_3}{2}$

TYPE B-2

XVI. When two of the zeros are pure imaginaries, the equation may be written:

$$\int \frac{dy}{y} = \frac{A_1 - a}{b_3(A_1^2 + A_2^2)} \int \frac{dx}{x - A_1} - \frac{A_1 - a}{2b_3(A_1^2 + A_2^2)} \int \frac{2x\,dx}{x^2 + A_2^2}$$
$$+ \frac{aA_1 + A_2^2}{b_3(A_1^2 + A_2^2)} \int \frac{dx}{x^2 + A_1^2}$$

Performing the integration, we have

$$\log y = \frac{A_1 - a}{b_3(A_1^2 + A_2^2)} \log (x - A_1) - \frac{A_1 - a}{2b_3(A_1^2 + A_2^2)} \log (x^2 + A_2^2)$$

$$+ \frac{aA_1 + A_2^2}{b_3 A_2 (A_1^2 + A_2^2)} \tan^{-1} \frac{x}{A_2} + \log y_0$$

Exponentiating, we have

$$y = \frac{y_0 (x - A_1)^{\frac{A_1 - a}{b_3(A_2^2 + A_1^2)}} e^{\frac{aA_1 + A_2^2}{b_3 A_2 (A_1^2 + A_2^2)} \tan^{-1} \frac{x}{A_2}}}{(x^2 + A_2^2)^{\frac{A_1 - a}{2b_3(A_1^2 - A_2^2)}}}$$

Let $b_3 (A_1^2 + A_2^2) = d$; $\frac{A_1 - a}{d} = 2m$; $aA_1 + A_2^2 = k$ Then

$$y = \frac{y_0 (x - A_1)^{2m} e^{\frac{k}{A_2 d} \tan^{-1} \frac{x}{A_2}}}{(x^2 + A_2^2)^m}$$

y_0 may be determined as in the previous case. The origin is at the mean.

TYPE B-3

XVII. If $F(x)$ is quadratic and the zeros are complex, the equation may be written

$$\int \frac{dy}{y} = \int \frac{(x-a)\,dx}{b_0+b_1 x+b_2 x^2} = \int \frac{(x-a)\,dx}{b_2\left[\left(x^2+\frac{b_1}{b_2}x+\frac{b_1^2}{4b_2^2}\right)+\left(\frac{b_0}{b_2}-\frac{b_1^2}{4b_2^2}\right)\right]}$$

Let $X = x+\frac{b_1}{2b_2}$; $A^2 = \frac{b_0}{b_2} - \frac{b_1^2}{4b^2}$

Then $x-a = X-\frac{b_1}{2b_2}-a = X+c$, where $c = -\left(\frac{b_1}{2b_2}+a\right)$

We have then

$$\log y = \int \frac{(X+c)\,dX}{b_2(X^2+A^2)} = \int \frac{X\,dX}{b_2(X^2+A^2)} + \frac{1}{b_2}\int \frac{c\,dX}{X^2+A^2}$$

$$= \frac{1}{2b_2}\log(X^2+A^2) + \frac{c}{Ab_2}\tan^{-1}\frac{X}{A} + \log y'$$

Exponentiating,

$$y = y'(X^2+A^2)^{\frac{1}{2b_2}}\, e^{\frac{c}{Ab_2}\tan^{-1}\frac{X}{A}}$$

$$= y_0\left(1+\frac{X^2}{A^2}\right)^{\frac{1}{2b_2}} e^{\frac{c}{Ab_2}\tan^{-1}\frac{X}{A}}$$

which may be written

$$y = y_0\left(1+\frac{X^2}{A^2}\right)^{-n} e^{-\gamma\tan^{-1}\frac{X}{A}},$$

where $n = -\frac{1}{2b_2}$; $-\gamma = \frac{c}{Ab_2}$; $A^2 = \frac{4b_0b_2-b_1^2}{4b_2^2}$

y_0 may be calculated as in Type A-1. Origin = Mean $-\frac{b_1}{2b_2}$

Type C-1

XVIII. When $F(x)$ is cubic, and two zeros of the denominator are equal, the equation may be written in the form

$$\int \frac{dy}{y} = \int \frac{(x-a)\,dx}{b_3(x-A_1)^2(x-A_2)} = \frac{A_1-a}{b_3(A_1-A_2)} \int \frac{dx}{(x-A_1)^2}$$

$$- \frac{A_2-a}{b_3(A_1-A_2)^2} \int \frac{dx}{x-A_1} + \frac{A_2-a}{b_3(A_1-A_2)^2} \int \frac{dx}{x-A_2}$$

Performing the integration, we have

$$\log y = \frac{-(A_1-a)}{b_3(A_1-A_2)(x-A_1)} - \frac{A_2-a}{b_3(A_1-A_2)^2} \log(x-A_1)$$

$$+ \frac{A_2-a}{b_3(A_1-A_2)^2} \log(x-A_2) + \log y'$$

Exponentiating,

$$y = \frac{y' e^{-\frac{A_1-a}{b_3(A_1-A_2)(x-A_1)}} (x-A_2)^{\frac{A_2-a}{b_3(A_1-A_2)^2}}}{(x-A_1)^{\frac{A_2-a}{b_3(A_1-A_2)^2}}} = \frac{y' e^{\frac{-m_1}{x-A_1}} (x-A_2)^{m_2}}{(x-A_1)^{m_2}}$$

where $m_1 = \dfrac{A_1-a}{b_3(A_1-A_2)}$; $m_2 = \dfrac{A_2-a}{b_3(A_1-A_2)^2}$

Now, changing the origin to A_1, i. e., replacing x by $x+A_1$, we have

$$y = \frac{y' e^{-\frac{m_1}{x}} (x+A_1-A_2)^{m_2}}{x^{m_2}}$$

$$= \frac{y_0 e^{-\frac{m_1}{x}} \left(1+x/k\right)^{m_2}}{x^{m_2}}$$

where $k = A_1 - A_2$

y_0 may be determined as in the previous case.
Origin = Mean + A_1.

TYPE C-2

XIX. If all the zeros are equal, the equation may be written

$$\int \frac{dy}{y} = \int \frac{A_1 - a}{b_3 (x - A_1)^3} dx + \int \frac{dx}{b_3 (x - A_1)^2}$$

Integrating,

$$\log y = -\frac{A_1 - a}{2b_3 (x - A_1)^2} - \frac{1}{b_3 (x - A_1)} + \log y_0$$

Exponentiating,

$$y = y_0 e^{\frac{-(A_1 - a)}{2b_3 (x - A_1)^2} - \frac{1}{b_3 (x - A_1)}} = y_0 e^{\frac{A_1 - a - 2x}{2b_3 (x - A_1)^2}}$$

where y_0 may be determined as before. Origin = Mean.

TYPE C-3

XX. When $F(x)$ is quadratic, and the zeros are real and equal, the equation may be written

$$\int \frac{dy}{y} = \int \frac{1}{b_2} \frac{(x - a)dx}{(x + \frac{b_1}{2b_2})^2}$$

$$= \int \frac{[(x + \frac{b_1}{2b_2}) - (a + \frac{b_1}{2b_2})] dx}{b_2 (x + \frac{b_1}{2b_2})^2}$$

$$= \int \frac{dx}{b_2 (x + \frac{b_1}{2b_2})} - \int \frac{(a + \frac{b_1}{2b_2})}{b_2 (x + \frac{b_1}{2b_2})^2} dx$$

$$\therefore \log y = \frac{1}{b_2} \log (x + \frac{b_1}{2b_2}) + \frac{a + \frac{b_1}{2b_2}}{b_2 (x + \frac{b_1}{2b_2})} + \log y'$$

Exponentiating,

$$y = y' \left(x + \frac{b_1}{2b_2}\right)^{\frac{1}{b_2}} e^{\frac{a + \frac{b_1}{2b_2}}{b_2\left(x + \frac{b_1}{2b_2}\right)}}$$

Now let

$$x + \frac{b_1}{2b_2} = x' \; ; \; \frac{1}{b_2} = -p \; ; \; \frac{a + \frac{b_1}{2b_2}}{b_2} = -\gamma$$

Then $y = y_0 \, x^{-p} e^{-\frac{\gamma}{x}}$

The constants in terms of moments are

$$-p = \frac{-4\mu_3}{\mu_3 + 2a\mu_2} \; ; \; -\gamma = \frac{4\mu_2(a\mu_3 - 2a^2\mu_2)}{\mu_3^2 - 4a\mu_2\mu_3 + 4a^2\mu_2}$$

It has been found[1] that $y_0 = \frac{N\gamma^{p-1}}{\Gamma(p-1)}$. Origin = Mean $- \frac{b_1}{2b_2}$

XXI. The following example, illustrated in the chart (p. 140' is given to illustrate the method. The data is fitted by Type A-7.

Ratio of Revenue to Net Worth in Traction Companies

Ratio	Observed Frequency	Theoretical Frequency
.04	7	7.3
.12	43	31.9
.20	48	55.6
.28	75	63.6
.36	53	57.6
.44	34	45.1
.52	25	32.2
.60	22	21.6
.68	12	14.1
.76	14	8.6
.84	5	5.2
.92	6	3.2
1.00	7	1.9

1. Elderton, "Frequency Curves and Correlation," p. 82.

The constants calculated from the observed frequencies were

$\mu_2 = 6.4354$	$b_0 = -5.7908$	$m = 24.97$
$\mu_3 = 18.0115$	$b_1 = -1.2125$	$p_1 = 5.3312$
Mean $= .376$	$b_2 = -.0334$	$p_2 = 35.2713$
Mode $= .2793$	$A_1 = 5.6587$	$\log y_0 = 50.1479$
$a = 1.2125$	$A_2 = 30.6287$	

Origin at Mean − 30.6287 or 30.4287 to left of 53 group.

Curve starts at $30.4287 - 24.97 = 5.4587$ before the center of this group.

EQUATION

$$y = y_0\,(x - 24.97)^{5.3312}\; x^{-35.2713}$$

Edwin D Mouzon, Jr.