NOTES

This section is devoted to brief research and expository articles, notes on methodology and other short items.

ORTHOGONAL POLYNOMIALS APPLIED TO LEAST SQUARE FITTING OF WEIGHTED OBSERVATIONS

BY BRADFORD F. KIMBALL

1. Introduction. Let the independent variable be denoted by x, and let it range over n consecutive integral values x_1 to x_n. Thus x represents the index-number of the ordered intervals at which observations are taken, where the intervals are all of equal length, and an index-number is assigned in consecutive order to every interval within the range of investigation, *whether observations occur in that interval or not.* Let y_x denote the observation measure (usually referred to as observed value), if such observation exists. Let w_x denote the weight of that observation, with weight zero assigned where observations are lacking.

To shorten the notation, summation over all values of x from x_1 to x_n will be denoted by the sign Σ. If a subscript and superscript is used, the context will indicate the variable to which the summation refers. The rth binomial coefficient will be denoted by $\binom{x}{r}$.

A system of polynomials $\phi_r(x)$, $r = 0, 1, 2, 3, \cdots$ of degree r in x is said to be an orthogonal system, for the purposes of this paper, if they satisfy the relations

$$\sum W_x \phi_r(x)\phi_s(x) \begin{cases} = 0, & r \neq s \\ \neq 0, & r = s. \end{cases} \tag{1}$$

To construct the polynomials, one may write them in the form

$$\begin{aligned} \phi_0(x) &= f_0(x) = \text{constant} \\ \phi_r(x) &= f_r(x) - \sum_{i=0}^{r-1} h_i \phi_i(x) \qquad r = 1, 2, 3, \cdots \end{aligned} \tag{2}$$

where the h_i are constants and the $f_r(x)$ are arbitrary polynomials of degree r. It then follows from the conditions of orthogonality that

$$h_i = \frac{\sum w_x f_r(x)\phi_i(x)}{\sum w_x [\phi_i(x)]^2}. \tag{3}$$

Thus when the polynomials $f_r(x)$ have been chosen for all r, the system of orthogonal polynomials for a given set of weights can be constructed and is uniquely determined except for a constant factor [1].

By virtue of the relation (2) and the conditions of orthogonality (1), it follows that

$$\Sigma w_x[\phi_r(x)]^2 = \Sigma w_x f_r(x)\phi_r(x). \tag{4}$$

Define the function $\Phi(r, k)$ by

$$\Phi(r, k) = \Sigma w_x f_r(x)\phi_k(x), \qquad r = 0, 1, 2, 3, \cdots . \tag{5}$$

It follows from the relations (2) and (3) that

$$\phi_r(x) = f_r(x) - \sum_{i=0}^{r-1} \frac{\Phi(r, i)}{\Phi(i, i)} \phi_i(x) \tag{6}$$

where it is to be noted that this summation is independent of x.

Define q_r and Y_r by

$$q_r = \Sigma w_x[\phi_r(x)]^2 = \Sigma w_x f_r(x)\phi_r(x) = \Phi(r, r), \tag{7}$$

$$Y_r = \Sigma w_x y_x \phi_r(x). \tag{8}$$

Then if $u_r(x)$ represents the polynomial solution of degree r of the normal equations set up for observed values y_x and weights w_x,

$$u_r(x) = \frac{Y_0}{q_0} + \frac{Y_1}{q_1}\phi_1(x) + \frac{Y_2}{q_2}\phi_2(x) + , \cdots , + \frac{Y_r}{q_r}\phi_r(x). \tag{9}$$

If E^2 denotes the weighted sum of the squares of the discrepancies between the ordinates $u_r(x)$ of the fitted curve and the observed values y_x, then [2],

$$E^2 = \sum w_x[u_r(x) - y_x]^2 = \sum w_x y_x^2 - \sum_{i=0}^{r} \frac{Y_i^2}{q_i}. \tag{10}$$

The practicability of the use of orthogonal polynomials is thus seen to depend upon whether the quantities $\Phi(r, k)$ and Y_r can be evaluated in a reasonably simple manner.

The thesis of this paper is that if $f_r(x)$ is taken as the binomial coefficient $\binom{x}{r}$, one can effectively apply the method of orthogonal polynomials. This is made possible by the use of factorial moments in conjunction with an adding machine that prints cumulative totals.

In treating the same problem Aitken sets up the normal equations in terms of factorials, but considers the explicit use of orthogonal polynomials impractical. He writes: "the arbitrary nature of the weights stands in the way of any analytical sophistication; orthogonal polynomials emerge, but are not of great use; and the necessity of solving the moment equations cannot be circumvented" [3]. He prefers a determinantal method of solution of the normal equations which the writer has found to be more involved from a practical point of view, than the present method, although it is elegant from a theoretical standpoint.

Thus although the present method is not new from the point of view of theory, the writer has found that forms made up by the use of the technique suggested below, offer an effective method for fitting polynomial curves to weighted observations.

2. Simplification of the problem when $f_r(x) = \binom{x}{r}$. Factorial moments S_r and M_r are defined by

$$S_r = \Sigma \binom{x}{r} w_x, \qquad M_r = \Sigma \binom{x}{r} w_x y_x \qquad r = 0, 1, 2, \cdots. \tag{11}$$

These moments are not difficult to compute and are readily checked as computed. Formula for $\Phi(r, k)$ then becomes

$$\Phi(r, k) = \Sigma \binom{x}{r} w_x \phi_k(x). \tag{12}$$

Thus since $\phi_0(x) = 1$, $\Phi(r, 0) = \Sigma \binom{x}{r} w_x = S_r$ and hence

$$\phi_1(x) = \binom{x}{1} - \frac{\Phi(1, 0)}{\Phi(0, 0)} = x - \frac{S_1}{S_0}.$$

Again

$$\begin{aligned}\Phi(r, 1) &= \Sigma \binom{x}{r} w_x \left(x - \frac{S_1}{S_0}\right) = \Sigma \binom{x}{r}\binom{x}{1} w_x - \frac{S_1}{S_0} \Sigma \binom{x}{r} w_x \\ &= (r + 1)S_{r+1} + rS_r - \frac{S_1 S_r}{S_0}.\end{aligned}$$

Hence

$$q_1 = \Phi(1, 1) = 2\,S_2 + \left(1 - \frac{S_1}{S_0}\right) S_1.$$

A recursion formula for $\Phi(r, k)$ may be obtained by expanding $\phi_k(x)$ in formula (12) by means of (6). Thus

$$\begin{aligned}\Phi(r, k) &= \sum \binom{x}{r}\binom{x}{k} w_x - \sum_{i=0}^{k-1} \frac{\Phi(k, i)}{q_i} \left[\sum \binom{x}{r} w_x \phi_i(x)\right] \\ &= \sum \binom{x}{r}\binom{x}{k} w_x - \sum_{i=0}^{k-1} \frac{\Phi(r, i)\Phi(k, i)}{q_i}.\end{aligned} \tag{13}$$

The first term can be easily expressed as a linear combination of binomial coefficients, and thus as a linear combination of moments S_i.

The formula for Y_r can be broken down as follows:

$$Y_0 = \sum w_x y_x = M_0,$$

(14)
$$Y_r = \sum w_x y_x \phi_r(x) = \sum w_x y_x \binom{x}{r} - \sum_{i=0}^{r-1} \frac{\Phi(r, i)}{q_i} [\sum w_x y_x \phi_i(x)]$$

$$= M_r - \sum_{i=0}^{r-1} \frac{\Phi(r, i)}{q_i} Y_i.$$

Thus

$$Y_1 = M_1 - \frac{S_1}{S_0} Y_0,$$

$$Y_2 = M_2 - \frac{\Phi(2, 1)}{q_1} Y_1 - \frac{\Phi(2, 0)}{q_0} Y_0, \quad \text{etc.}$$

3. General technique of computation. In determining the best fitting polynomial of degree r, the ratios $\Phi(r, i)/q_i$ are seen to play an important part. In a form for calculation, these quantities should receive simple designations such as b_i for a second degree curve, c_i for a third degree curve, etc. Suppose they are designated by R_i for a curve of degree r; then

$$\phi_r(x) = \binom{x}{r} - \sum_{i=0}^{r-1} R_i \phi_i(x) \tag{15}$$

$$Y_r = M_r - \sum_{i=0}^{r-1} R_i Y_i \tag{16}$$

$$q_r = \sum \binom{x}{r}^2 w_x - \sum_{i=0}^{r-1} R_i \Phi(r, i) \tag{17}$$

and in determining $\Phi(r, k)$ for $k = 0, 1, 2, \cdots r - 1$, formula (13) may be written:

$$\Phi(r, k) = \sum \binom{x}{r}\binom{x}{k} w_x - \sum_{i=0}^{k-1} R_i \Phi(k, i). \tag{18}$$

The fact that these quantities R_i appear as multipliers in so many of the fundamental formulas greatly simplifies the mechanics of the calculation, especially when a calculating machine is used.

In final determination of polynomial curve the differences of the polynomial at $x = 0$ are readily determined since the leading term of each orthogonal polynomial is a binomial coefficient and thus

(19)
$$\Delta^k \phi_r(0) = -\sum_{i=0}^{r-1} R_i \Delta^k \phi_i(0), \qquad k = 1, 2, 3, \cdots, r - 1$$

$$\Delta^r \phi_r(0) = 1.$$

Since the effectiveness of the method depends upon the availability of an adding machine which records a cumulative subtotal, the determination of the curve from the differences at the point $x = 0$ is not a hardship and indeed affords a quick and accurate means of setting up the curve for purposes of plotting and checking.

$$\begin{aligned} u_r(0) &= \frac{Y_0}{q_0} + \frac{Y_1}{q_1}\phi_1(0) + \frac{Y_2}{q_2}\phi_2(0) + , \cdots , + \frac{Y_r}{q_r}\phi_r(0), \\ \Delta^k u_r(0) &= \frac{Y_k}{q_k} + \frac{Y_{k+1}}{q_{k+1}}\Delta^k \phi_{k+1} + , \cdots , + \frac{Y_r}{q_r}\Delta^k \phi_r(0), \\ \Delta^r u_r(0) &= \frac{Y_r}{q_r}. \end{aligned} \tag{20}$$

The advantage of the use of orthogonal polynomials becomes particularly apparent when error formulae are to be used. The formula for the sum of the squares of the discrepancies, denoted by E^2, is given above (formula (10)). The estimated variance V of the weighted observations about the fitted curve is thus $E^2/(n - r - 1)$ where n is the number of values of x used in fitting and r is the degree of the curve fitted. Recalling that the matrix of the normal equations is of the diagonal form with diagonal elements $q_0, q_1, \cdots, q_r$ it follows that the coefficient Y_k/q_k of $\phi_k(x)$ in the expansion of $u_r(x)$ has the variance V/q_k.

Furthermore the variance of the ordinate of the fitted curve $u_r(x)$ at a point x due to sampling variations in the determination of the coefficients of the curve, under the assumption that the weights and values of the independent variable x do not involve errors, has the simple form

$$\begin{matrix}\text{Variance of } u_r(x) \\ \text{at point } x\end{matrix} = V\left[\frac{\phi_0^2(x)}{q_0} + \frac{\phi_1^2(x)}{q_1} + , \cdots , + \frac{\phi_r^2(x)}{q_r}\right] \tag{21}$$

since the covariances of the orthogonal polynomials are zero [4].

REFERENCES

[1] G. Szego, *Orthogonal Polynomials*, Colloquium Publications, Amer. Math. Soc., New York, 1939, Vol. 23, Chapter II.

[2] Max Sasuly, *Trend Analysis of Statistics*, The Brookings Institution, Washington, 1934, pp. 296–297.

[3] A. C. Aitken, "On Fitting Polynomials to Weighted Data by Least Squares," *Proc. Roy. Soc. of Edinburgh*, Session 1933–34, Vol. 54, Part I, No. 1, p. 2.

[4] Henry Schultz, "The Standard Error of a Forecast from a Curve," *Jour. Amer. Stat. Assn.*, June 1930; or Whittaker & Robinson, *The Calculus of Observations*, Blackie & Son Ltd., London, 1937, 2d ed., p. 242.

Port Washington, N. Y.