SASULY ON STATISTICS

Trend Analysis of Statistics—Theory and Technique. By Max Sasuly. Washington, D.C., The Brookings Institution, 1934. xiii+421 pp.

The formal procedure for fitting polynomial curves by the method of least squares has been available, from the investigations of Legendre, Gauss, and Laplace, since the early part of the nineteenth century. However, its utility was, until lately, seriously limited by the excessive labor entailed in carrying through the numerical computations even for polynomial curves of relatively low degree. The most fruitful research in recent years has been directed to the development of an effective technique to simplify the numerical computation steps in cases in which the data are equally spaced and equally weighted. As a result of these investigations, the practical difficulties which were formerly encountered in these cases have been entirely removed.

The factors which have contributed mainly to the notable advance which has thus lately been achieved in facilitating the practical least squares fitting of polynomial curves are the use of the orthogonal form of polynomials in the fitting procedure, together with the employment of the Gregory-Newton expansions for these orthogonal polynomials. The advantages of the orthogonal representation had long been recognized and, in fact, the foundations of the theory of least squares orthogonal polynomial fitting had been laid by Tchebycheff in the middle of the last century. However, the practical application of his methods was very laborious. Then, in 1913, the practical least squares fitting of polynomial curves was greatly facilitated by the proposal of Sheppard to use the Gregory-Newton form of expansion for the equation of the polynomial curve to be fitted, that is, an expansion in a series of factorial terms, $1, x, x(x-1)/2!, x(x-1)(x-2)/3!, \cdots$, rather than in the usual form of a series of power terms, $1, x, x^2, x^3, \cdots$, owing to the fact that an explicit solution for the parameters of a polynomial of any degree could then be secured. Moreover, by using that form of expansion, the values of the parameters were obtained in terms of factorial moments of the data, which could readily be computed by a repeated summation procedure. The final stage in simplifying the practical least squares fitting of polynomial curves consisted in combining the advantages secured from the orthogonal representation with those resulting from the use of the factorial form of expansion of the polynomials.

In the book under review, the author has developed the theoretical basis of the underlying formulas for fitting polynomial curves by the method of least squares in much the same order as that which was actually followed in the course of the historical development, as briefly indicated above. That is, the formulas for least squares fitting of polynomial curves for a few of the lower degree polynomials in the familiar power series form are first derived. Up to the present time, however, no direct general solution for the values of the parameters of least squares power polynomials in terms of the power moments of the data appears to be available. Thus, attention is directed to the Gregory-Newton form of expansion, where a relatively simple, general solution in terms of factorial moments of the data is obtained. The procedure for computing the factorial moments by means of repeated summations of the data is indicated. Finally, consideration is given to the least squares fitting of polynomial curves by the use of orthogonal systems of Gregory-Newton polynomials.

In addition to the development of methods for fitting polynomial curves to series of data by the method of least squares, the book also treats the smoothing of a series of data by means of moving polynomial arcs, and by functions of a somewhat broader character which may be derived from them.

In brief, then, it may be said that the purpose of this book is to exhibit the possibilities of trend representation by least squares polynomials, and of the smoothing, or graduation, of series of data by means of moving least squares polynomial arcs and closely related functions. The author states that the primary aim of this book is to "derive formulas and computation schedules that will simplify as much as possible effective, practical trend analysis . . . The theoretical basis of the underlying formulas is treated to an extent necessary to exhibit their derivation, and to disclose the simple relations between the new formulas introduced and others more familiar in the literature. This gives the necessary background as well as the full scope of the new method presented. However, the formulas and fitting procedure that are of primary interest to the computer may be applied independently of the formal proofs of their derivations . . . The procedure steps given in the illustrative examples . . . are given as nearly as possible in the order used in actual practice, and in sufficient detail so that their principal applications may be made without reference to the text discussion."

With this orientation concerning the general scope and aim of the book, it will perhaps be possible to proceed with clearer understanding to a detailed consideration of the contents.

The text is divided into three parts, or books. The title of Book I is *Fitted polynomials in trend analysis*. The first chapter in this book, Chapter 2, is devoted to general considerations concerning the fitting properties of polynomials, followed by a brief discussion of various types of fitting criteria, and closes with a general treatment of the problem of fitting functions by the criteria of least squares. The next chapter takes up the least squares fitting of mid-range power polynomials, and derives explicit formulas for the parameters in terms of the mid-range power moments of the data for polynomials through the sixth degree. A method for building up the orthogonal form of the mid-range power polynomials is indicated in the latter part of this chapter, and the explicit expression for the first ten of these polynomials is given. Chapter 4 deals with mid-range parameters and fitted values as weighted data averages. Here the points on the mid-range power polynomials are expressed explicitly in terms of the data to which the curve is fitted, and the values of the parameters are derived in terms of the data differences. The expressions for these higher order parameters can be used to determine certain important characteristics of the fitted curve, such, for example, as the slope and the curvature at various parts of the range. In Chapter 5, formulas are derived for interpolation by fitted mid-range power polynomials for any spacing interval. Chapter 6 discusses general polynomial interpolation by means of the Lagrange formula and the Newton divided difference formula, which are of use when the data are unequally spaced. In Chapter 7, the last chapter in Book I, attention is directed to central difference interpolation by the Newton-Stirling, Newton-Bessel, and Everett formulas, and to smooth junction interpolation by means of the Karup cubic and Glover quintic formulas. The subjects of inverse interpolation and of approximate integration and summation are also treated in this chapter. In brief, then, Book I may be said to deal mainly with the least squares fitting of mid-range power polynomials, and with various methods of interpolation. For the most part, the results are in the form most familiar in the literature on these subjects.

The subject matter of Book II is *Trend determination by data sums and averages*. Chapter 8, the first chapter in Book II, deals with power moments and factorial moments of the data. The method for computing factorial moments by repeated forward summations of the data, as well as by repeated backward summations, is developed. The relations between factorial moments and power moments about the same origin, and also about arbitrary origins, are here derived. The chapter concludes with a discussion of moments and averages, in which formulas are derived for the values of certain useful statistical data averages in terms of factorial moments. Sheppard's moment corrections are also discussed here. Chapters 9 and 10 discuss the graduation of a series of data by means of moving least squares polynomial arc mid-values and iterated moving average functions. First, the procedure for computing the mid-values of moving least squares parabolic, quartic, and sextic arcs in terms of repeated summations of the data is indicated. However, the computation of moving arc mid-points becomes quite laborious for the higher degree arcs, and a method is indicated for building up various synthetic arc formulas which will secure approximately the same degree of smoothing as the moving sextic arc mid-value, but which is much easier to compute than the latter. These synthetic arc formulas involve the use of iterated moving averages, and a method is indicated for the computation of iterated moving average functions in terms of repeated summations of the data. Consideration is also given to the uses which can be made of graduation formulas in smoothing time series data. The remaining chapters in Book II deal with moment equivalence conditions for function arcs, and function parameters in terms of factorial moments.

Analytic trend fitting by general systems of polynomials is the subject matter of Book III. In the first chapter, Chapter 13, the useful Gregory-Newton form of expansion is employed to obtain the explicit expressions for the parameters of the least squares polynomial of any degree. Chapters 14 and 15 deal with orthogonal systems of Gregory-Newton polynomials. The method for building up such a system is indicated, the advantages of this form of representation in practical least squares polynomial curve fitting are set forth, and a convenient fitting procedure is developed. Chapter 16 takes up special polynomial trend fitting in integral form, and the last chapter in the book discusses interpolation and sub-tabulation by fitted polynomials. A group of tables for use in connection with the various procedures developed in the text are inserted at the end of the book.

This book should be of value both to the computer whose primary interest is in the acquisition of effective, practical, curve fitting methods, as well as to those who may wish to follow the theoretical development of the effective methods which have lately been devised for fitting polynomial curves by the method of least squares. From the standpoint of this latter class, a very desirable feature of the book is the large number of references to source material.

It is to be regretted that no reference is made to three important papers on the least squares fitting of orthogonal polynomials which have recently appeared. It is possible, of course, that the latest of these papers had not been published at the time the manuscript of this book was completed. In the order of their appearance, these three papers are: (1) *On graduation according to the method of least squares by means of certain polynomials*, by Fredrik Esscher (Försäkringsaktiebolaget Skandia, 1855–1930, Stockholm, vol. 2 (1930), p. 107); (2) *Approximation and graduation according to the principle of least squares by orthogonal polynomials*, by Charles Jordan (Annals of Mathematical Statistics, vol. 3 (1932), p. 257); and (3) *On the graduation of data by the orthogonal polynomials of least squares*, by A. C. Aitken (Proceedings of the Royal Society of Edinburgh, vol. 53 (1933), p. 54). Esscher's paper is noteworthy primarily for its extremely simple development of the general form for the least squares orthogonal (Tchebycheff) polynomials. Esscher's development is very similar to that used by Goursat* in deriving the Legendre polynomials, to which the Tchebycheff polynomials are closely related. Jordan has published a number of papers on the subject of least squares orthogonal polynomial curve fitting, to some of which reference has been made in the book under review. His 1932 paper, which is mentioned above, may be said to round out his earlier work. In this paper he has appended some excellent bibliographical and historical notes which deal with the results obtained by other investigators in this field. Aitken's paper is outstanding not only for its simple and elegant development of the form of the Tchebycheff polynomials, which is similar to that given by Esscher, but also for presentation of extremely simple fitting procedure.

Throughout the book under review, the material is treated in a manner which makes few demands on the mathematical ability of the reader. In some instances, the method of development which is followed has resulted in the inclusion of perhaps an undue amount of detail. For example, in Chapter 8 it seems to us that the relationships between power moments and factorial moments, and between moments taken about different origins, could have been developed in a more compact and simple manner by making use of Vandermonde's theorem, together with the so-called differences of zero and differential coefficients of zero, tables of which are given later in Chapter 17, than by means of the somewhat laborious and extended algebraic procedure which was actually used. Likewise, the method which has been followed in resolving the least squares Gregory-Newton polynomials into orthogonal components in Chapters 14 and 15 is decidedly more involved and difficult, in our opinion, than the derivation of the orthogonal polynomials which is given in the above-mentioned papers of Esscher and Aitken.

The book contains a number of typographical errors. Fortunately, most of them are of such nature as to be readily apparent to the reader.

K. W. HALBERT

* Goursat-Hedrick, *Mathematical Analysis*, 1904, vol. 1, p. 173.