APPLICATION OF ORTHOGONAL POLYNOMIALS IN FITTING AN ASYMPTOTIC REGRESSION CURVE

By A. S. Chopra *

Institute of Agricultural Research Statistics, New Delhi

I. Introduction

The asymptotic regression curve $Y = \alpha - \beta\rho^x$ is widely used in describing non-linear regression. This functional form has been found to fit particularly well in describing yield-dose relationship with fertilizer application to agricultural crops.

In estimating the parameters α, β and ρ of the asymptotic regression curve, ρ presents the major difficulty when more than three equally spaced values of x's are given (the value of x may be taken as 1, $2 \cdots n$). If ρ were known, α and β can be obtained by fitting a linear regression of Y and ρ^x. Various methods of estimating ρ have been proposed by Hartley (1948), Stevens (1951), Gomes (1953) and Patterson (1956). The limitation of these methods are many, for example Stevens (1951) method involves the inversion of the information matrix and the tables are available only for $n = 5$, 6 and 7 for $P = 0 \cdot 25$ to $0 \cdot 70$ and that of Gomes (1953) involves in finding out a root of the equation.

$$P_1(Z) \Sigma Y + P_2(Z) \Sigma xYZ^m + P_3(Z) \Sigma YZ^m = 0$$

where $P_1(Z)$, $P(Z_2)$ and $P_3(Z)$ are polynomials in Z and Z lies between 0 and 1. If the equation has no root between 0 and 1, the fitting of the curve cannot be carried out. On the other hand, if two or more such roots were to be found it would not be possible to choose among them and the tables for the solution are available only for $n = 4$ or 5. Patterson (1956) has proposed an ingenious and simple method of estimating ρ as the ratio of two contrasts. The coefficients of the contrasts were obtained with the help of tables prepared by Stevens (1951). He showed that this estimate was highly efficient. The method proposed by Patterson wholly depends on Stevens' method which is difficult for n greater than seven. Thus Patterson could give estimate for ρ for $n = 4$, 5, 6 and 7 and they are fully efficient for $n = 4$ when $\rho = 0 \cdot 26$ and $n = 5$, 6 and 7 when ρ lies in the range $0 \cdot 3$ to $0 \cdot 5$.

* Present address: Directorate of Flight Safety, Air Headquarters, New Delhi.

Shah (1961) extended Patterson's method to the problem of fitting the curve $Y = \alpha + \beta x + \lambda \rho^x$ for equally spaced ordinates.

In the present discussion, a procedure is given for fitting the extended asymptotic regression curve of the type of $Y = a_0 + a_1 x + \cdots a_k x^k + \lambda \rho^x$ for equally spaced levels of x and the results of asymptotic regression curve $Y = \alpha - \beta \rho^x$ have been compared with those of Patterson's method.

2. The New Approach

Consider $Y = a_0 + a_1 x + \cdots + a_k x^k + \lambda \rho^x$.

This curve involves $k+3$ number of parameters, *viz.*, $a_0, a_1 \cdots a_k$, λ and ρ. Of all the parameters ρ presents the major difficulty in estimation. Once the estimate of ρ is known, the other parameters are obtained by fitting a multiple linear regression of Y on x, $x^2, \cdots x^k$ and ρ^x.

ρ has been expressed as the ratio of two contrasts where the coefficients of the contrast corresponds to the coefficients of ξ'_{k+1} corresponding to $n-1$ when n is the number of levels. If the number of levels are greater than the number of parameters then this will serve as an approximate estimate to ρ.

$$\rho = \frac{\mu_1 y_n + \mu_2 y_{n-1} + \cdots + \mu_{n-1} y_2}{\mu_1 y_{n-1} + \mu_2 y_{n-2} + \cdots + \mu_{n-1} y_1} = \frac{A}{B}$$

where

(i) $\sum_{i=1}^{n-1} \mu_i = 0$ and coefficients are those of ξ'_{k+1} corresponding to $n-1$.

(ii) $E(Y_i) = a_0 + a_{1.}\, i + a_2 i^2 + \cdots + a_k i^2 + \lambda \rho^i$.

(iii) $E(Y_i - y_i)(Y_j - y_j) = \sigma^2$ if $i = j$,
$= 0$, if $i \neq j$.

$$E\left(\frac{A}{B}\right) \simeq \frac{E(A)}{E(B)} \text{ by neglecting bias.}$$

$$E(A) = -\beta \rho^2 (\mu_1 \rho^{n-2} + \cdots + \mu_{n-1})$$

$$E(B) = -\beta \rho (\mu_1 \rho^{n-2} + \cdots + \mu_{n-1})$$

$$E\left(\frac{A}{B}\right) = \rho$$

$$V\left(\frac{A}{B}\right) = \left\{ E\left(\frac{A}{B}\right)\right\}^2 \left\{\frac{\text{Var}(A)}{(E(A))^2} + \frac{\text{Var}(B)}{(E(B))^2} - \frac{2\,\text{Cov}(A, B)}{E(A)\,E(B)}\right\}$$

$$= \frac{\phi^2}{\rho^2} \frac{\left\{(1+\rho^2)\sum_{i=1}^{n-1} \mu_i^2 - 2\rho \sum_{i<j} \mu_i \mu_j\right\}}{(\mu_1 \rho^{n-2} + \cdots + \mu_{n-1})^2}$$

where $\phi^2 = \sigma^2/\beta^2$.

3. Illustration

Suppose the regression curve to the fitted is $Y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \lambda\rho^x$ when the number of levels are either 6 or 7. For $n = 6$, the exact value of ρ is given by the ratio of two quartic ξ_4' contrasts corresponding to $n - 1 (= 5)$ whose coefficients are $\mu_1 = 1$, $\mu_2 = -4$, $\mu_3 = 6$, $\mu_4 = -4$, $\mu_5 = 1$. The estimate of ρ is given by

$$\rho = \frac{Y_6 - 4Y_5 + 6Y_4 - 4Y_3 + Y_2}{Y_5 - 4Y_4 + 6Y_3 - 4Y_2 + Y_1}.$$

For $n = 7$, the estimate of ρ is given by the ratio of two quartic contrasts corresponding to $n - 2 (= 6)$ and whose coefficients are $\mu_1 = 1$, $\mu_2 = -3$, $\mu_3 = 2$, $\mu_4 = 2$, $\mu_5 = -3$, $\mu_6 = 1$ and the estimate is given by:

$$\frac{Y_7 - 3Y_6 + 2Y_5 + 2Y_4 - 3Y_3 + Y_2}{Y_6 - 3Y_5 + 2Y_4 + 2Y_3 - 3Y_2 + Y_1}.$$

4.1. Relative efficiency of the different methods.

Fitting of $Y = \alpha - \beta\rho^x$.

For the fitting of this curve the following values of μ were proposed by Patterson for different values of n:

$n = 4$, $\mu_1 = 4$, $\mu_2 = 1$, $\mu_3 = -5$

$n = 5$, $\mu_1 = 4$, $\mu_2 = 3$, $\mu_3 = -1$, $\mu_4 = -6$

$n = 6$, $\mu_1 = 4$, $\mu_2 = 4$, $\mu_3 = 2$, $\mu_4 = -3$, $\mu_5 = -7$

$n = 7$, $\mu_1 = 1$, $\mu_2 = 1$, $\mu_3 = 1$, $\mu_4 = 0$, $\mu_5 = -1$, $\mu_6 = -2$.

If the values of μ are taken as the coefficient of ξ_1' corresponding to $n - 1$, we get an estimate of ρ.

$n = 4$, $\mu_1 = 1$, $\mu_2 = 0$, $\mu_3 = -1$

$n = 5$, $\mu_1 = 3$, $\mu_2 = 1$, $\mu_3 = -1$, $\mu_4 = -3$

$n = 6$, $\mu_1 = 2$, $\mu_2 = 1$, $\mu_3 = 0$, $\mu_4 = -1$, $\mu_5 = -2$

$n = 7$, $\mu_1 = 5$, $\mu_2 = 3$, $\mu_3 = 1$, $\mu_4 = -1$, $\mu_5 = -3$, $\mu_6 = -5$

and for higher values of n, refer to Fisher and Yates tables.

The percentage efficiencies for $n = 4, 5, 6$ and 7 have been calculated from large sample formula for the variance of ρ and the most efficient estimate and are presented in Table I.

TABLE I

Percentage efficiencies of proposed estimates of ρ

ρ \ n	4	5	6	7
0·1	83·6 (95·4)	71·2 (88·5)	61·8 (79·2)	54·5 (74·5)
0·2	89·4 (98·6)	80·0 (95·7)	72·1 (90·2)	65·4 (86·8)
0·3	93·3 (99·8)	86·3 (99·2)	79·9 (97·0)	74·5 (95·0)
0·4	95·9 (99·0)	90·3 (99·9)	85·2 (99·6)	80·6 (98·1)
0·5	97·7 (97·4)	92·7 (98·9)	87·4 (98·8)	83·8 (96·8)
0·6	98·7 (98·5)	94·1 (97·0)	89·1 (96·1)	84·7 (92·4)
0·7	99·3 (97·4)	94·8 (94·7)	89·4 (92·5)	84·6 (86·9)
0·8	99·7 (96·3)	95·1 (92·3)	89·4 (88·8)	83·9 (81·3)
0·9	99·97 (95·3)	95·4 (90·0)	89·3 (85·2)	83·6 (76·3)

Note.—The figures in brackets are the efficiencies by Patterson's method.

The rate of decrease in efficiency with the increase in the value of n will be small, for instant, for $\rho = 0·5$ efficiency for $n = 8$ and 9 is 80·2 and 77·8 in comparison with the most efficient estimate. The minimum efficiency for $\rho = 0·5$ for higher value will be more than 63%. Thus the efficiency of the proposed estimate relative to Patterson's estimate can be obtained by dividing the percentage efficiency of proposed estimate by the percentage efficiency of Patterson's method.

4·2. Fitting of the regression curve

$$Y = \alpha + \delta x + \beta \rho^x.$$

The following μ coefficients for $n = 5, 6, 7$ and 8 were proposed by Shah (1961) for the estimation of ρ

$n = 5, \mu_1 = 5, \mu_2 = -4, \mu_3 = -7, \mu_4 = 6$

$n = 6, \mu_1 = 10, \mu_2 = -1, \mu_3 = -13, \mu_4 = -11, \mu_5 = 15$

$n = 7, \mu_1 = 10, \mu_2 = 1, \mu_3 = -6, \mu_4 = -14, \mu_5 = -8, \mu_6 = 17$

$n = 8, \mu_1 = 10, \mu_2 = 5, \mu_3 = -5, \mu_4 = -12, \mu_5 = -11,$

$\mu_6 = -7, \mu_7 = 20.$

The μ coefficients for the proposed estimates are

$n = 5, \mu_1 = 1, \mu_2 = -1, \mu_3 = -1, \mu_4 = 1$

$n = 6, \mu_1 = 2, \mu_2 = -1, \mu_3 = -2, \mu_4 = -1, \mu_5 = 2$

$n = 7, \mu_1 = 5, \mu_2 = -1, \mu_3 = -4, \mu_4 = -4, \mu_5 = -1, \mu_6 = 5$

$n = 8, \mu_1 = 5, \mu_2 = 0, \mu_3 = -3, \mu_4 = -4, \mu_5 = -3, \mu_6 = 0,$

$\mu_7 = 5$, etc.

The percentage efficiencies for $n = 5$, 6, 7 and 8 have been calculated from large sample formulae for the variance of ρ and the efficient estimate and are presented in Table II.

TABLE II

Percentage efficiencies of proposed estimate of ρ

ρ \ n	5	6	7	8
0·1	89·2 (98·4)	80·3 (97·3)	72·8 (96·0)	66·6 (94·4)
0·2	92·9 (99·7)	86·4 (99·7)	80·8 (99·0)	75·7 (97·8)
0·3	95·5 (99·7)	90·7 (99·7)	86·3 (98·4)	82·3 (96·9)
0·4	97·2 (99·6)	93·6 (98·4)	89·8 (95·5)	86·4 (93·0)
0·5	98·4 (99·0)	95·1 (96·1)	91·6 (91·4)	88·4 (87·5)
0·6	99·0 (98·1)	96·1 (93·7)	92·6 (87·1)	90·1 (82·6)
0·7	99·6 (97·3)	96·9 (91·4)	93·0 (83·0)	89·0 (76·1)
0·8	99·8 (96·3)	97·6 (89·4)	92·9 (79·0)	89·3 (71·8)
0·9	99·7 (95·4)	97·2 (86·7)	93·1 (75·9)	89·6 (68·2)

Note.—Figures in brackets correspond to efficiency by Shah's method.

5. *Example.*—As an example of the procedure when $n = 7$. The Y are average coefficients of correlations between yields separated by different periods.

Y ..	1·0000	0·7477	0·6851	0·6817	0·5330	0·5586	0·4919
x ..	0	1	2	3	4	5	6

The estimate of ρ is given by

$$r = \frac{3Y_6 + 2Y_5 + Y_4 - Y_3 - 2Y_2 - 3Y_1}{3Y_5 + 2Y_4 + Y_3 - Y_2 - 2Y_1 - 3Y_0} = 0\cdot 6654.$$

values of r^x are

r^x:	1·0000	0·6654	0·4428	0·2746	0·1960	0·1304	·0868

The regression of Y on r^x is

$$\frac{\Sigma (Y - \bar{Y})(r^x - \bar{r}^x)}{\Sigma (r^x - \bar{r}^x)^2} = 0\cdot 5058.$$

This is an estimate of $-\beta$ and α is estimated by

$$\hat{a} = \bar{Y} + \beta\bar{r}^x = 0\cdot 4677.$$

The regression equation becomes

$$Y = 0\cdot 4677 + 0\cdot 5058\,(0\cdot 6654)^x.$$

Variance of the estimate of $\rho = 13\cdot 2092\,\sigma^2$.

By Patterson's method, the equation becomes

$$Y = 0\cdot 4952 + 0\cdot 4884\,(0\cdot 6169)^x.$$

Variance of the estimate of $\rho = 14\cdot 5874\,\sigma^2$. Percentage efficiency of the proposed estimate $= 110\cdot 43$.

The predication power of the equation will be measured by the simple correlation r_{YYe} where Y and Ye are observed and expected values respectively.

i.e.,
$$r_{YYe} = \sqrt{1 - \frac{\Sigma (Y - Ye)^2}{\Sigma Y^2}}$$

and this will approach one as $Y - Ye$ tends to zero.

6. Summary

A method of getting an estimate of ρ in the non-linear regression curve $Y = a_0 + a_1 x \cdots + a_k x^k + \lambda\rho^x$ have been suggested with the

help of orthogonal polynomial coefficients and two particular cases when all a_i's are zero except a_0 and a_0 & a_1 are compared with their available methods.

ACKNOWLEDGEMENT

The author wishes to express his thanks to Shri T. P. Abraham, Assistant Statistical Adviser, I.C.A.R., for his helpful suggestions and critically examining the manuscripts.

REFERENCES

1. Finney, D. J. .. "The efficiencies of the alternative estimator for an asymptotic equation," *Biometrika*, 1958, **45.**
2. Fisher, R. A. and Yates, F. *Statistical Tables for Biological, Agricultural and Medical Research*, Oliver & Boyd, London, 1957.
3. Gomes, P. F. .. "The use of Mitschertich's regression law in the analysis of experiments with fertilizers," *Biometrics*, 1953, 9.
4. Hartley, P. O. .. "The estimation of non-linear parameters by internal least squares," *Biometrika*, 1948, **35.**
5. Nair, K. R. .. "The fitting of growth curves," *Statistics and Mathematics in Biology*, by Kempthorne, Lush and Bancraft.
6. Patterson, H. D. .. "A simple method for fitting an asymptotic curve," *Biometrics*, 1956, **12.**
7. Shah, B. K. .. "A simple method of fitting the regression curve $Y = \alpha + \delta x + \beta\rho^x$," *Ibid.*, 1961, **17.**
8. Stevens, W. L. .. "Asymptotic regression," *Ibid.*, 1951, **7.**
9. ———— .. *Report on Standardisation of Horticultural Experiments, I.A.R.S. (I.C.A.R.)*, Unpublished.