CHAPTER VI

METHOD OF FITTING THE SOME DOUBLE EXPONENTIAL REGRESSION

Tootill [38, 39, 40,] has described the computational procedure for the model

$$Z = \alpha + \beta \phi (\gamma, y)$$

where $\phi(\gamma, y)$ is any non-linear function of y involving a single unknown parameter γ.

Many situations in biology, epidemiology, economics and the physical sciences yield data which can best be described by linear combinations. This is particularly true in biological radiation, growth and tracer studies.

In this chapter we will discuss the more general and complex model containing two unknown parameters of the type :

$$Z = \alpha + \beta \phi (\gamma_1, \gamma_2, y) \quad \text{and}$$

$$Z = \alpha + \beta_1 x + \beta_2 \phi (\gamma_1, \gamma_2, y) .$$

In particular, we propose the following two new curves;

(i) $y = A + B_1 d_1^x + B_2 d_2^x$ and

(ii) $y = A + Hx + B_1 d_1^x + B_2 d_2^x$.

Note that if $d_2 = 0$ or $B_2 = 0$, we get the Makeham's modified Gomportz's curve, i.e. containing single non-linearity.

This curves adequately fit the data on force of mortality μ_x or colog (probability of survivors p_x) . While fitting the curves, Indira Bhanot [14] has used the weighting coefficients proportional to $l_x / \mu_x(1-\mu_x)$ $\{$ or $p_x l_x / (1-p_x) \}$, where l_x is the no. of persons exposed to risk. Perhaps we can approximately use the weights as proportional to l_x. We have fitted one of the curves to the Makeham's data [35] from ages 3 to 27 and observed that it fitted well and no other known curve fitted so nicely to the same data, Khatri and Shah [16].

Recently S.Lipton and C.McGilchrist [21] has described the technique for obtaining maximum likelihood estimates for the parameters using a combination of Steven's and Richard's [30] methods applied to the double exponential case. Richard G. Cornell [3,4] has shown how the method of partial totals may be used to fit linear combinations of any number of exponentials to data taken at equally spaced intervals. This method has been applied to special cases of the general model before, for instance, by Stevens [37] and Croxton and Cowden [5]. This method is often useful in computing preliminary estimates for iterative maximum likelihood solution i.e. if computing facilities are not available for obtaining maximum likelihood estimates, the method provides a systematic, consistent and relatively simple estimation procedure.

In this chapter we will describe the method known as internal least squares given by H. O. Hartley [10] and Indira Bhanot [14]. This method can be extended to any number of exponentials. In particular, the internal regressions of (i) and (ii) curves are obtained by direct summing the difference equations :

$$\text{(iii)}\quad y_{i+2}-(d_1+d_2)\, y_{i+1}+d_1d_2y_i = A(1-d_1)\,(1-d_2) \quad \text{and}$$

$$\text{(iv)}\quad y_{i+2}-(d_1+d_2)\, y_{i+1}+d_1d_2y_i = A(1-d_1)\,(1-d_2) + H(2-d_1-d_2) + Hi\,(1-d_1)\,(1-d_2).$$

When the observations are equally spaced, the different consequences of internal regressions due to (iii) and (iv) are also considered.

2. Modified method of Internal least squares:

2.1 To fit $y = A + B_1d_1^x + B_2d_2^x$.

The difference equation (iii) can be rewritten as

$$(e_{i+2}-e_{i+1}) - (e_{i+1}-e_i) = r_2 e_{i+2} - r_1 e_{i+1} + 2a \quad \ldots \quad (1)$$

where $d_1+d_2 = (2-r_1)/(1-r_2)$,

$d_1 d_2 = 1/(1-r_2)$,

$e_i = y_i - Z$, Z being any value, and

$A = Z - 2a/(r_2-r_1)$.

Case 1:- Let x takes values 0,1,2,....,n.

Then summing (1) over the values of i from 0 to i-2, we have,

$$\sum_{i=0}^{i-2} (e_{i+2} - e_{i+1}) - \sum_{i=0}^{i-2} (e_{i+1}-e_i) = r_2 \sum_{i=0}^{i-2} e_{i+2} - r_1 \sum_{i=0}^{i-2} e_{i+1}$$

$$+2(i-1)a.$$

$$\text{i.e. } e_i - e_{i-1} = r_2(e_2+e_3+\ldots e_i) - r_1(e_1+e_2+\ldots+e_{i-1}) + 2(i-1)a + e_1 - e_0$$

$$= r_2(e_0+e_1+\ldots+e_i) - r_1(e_0+e_1+\ldots+e_i) + r_1 e_0 + r_1 e_i$$

$$- r_2(e_0+e_1) + r(i-1)a + e_1 - e_0 .$$

$$\text{i.e. } e_i - e_{i-1} = rS_{1.i} + r_1 e_i + 2a(i-1) + b' \quad \ldots \quad (2)$$

where $r = r_2 - r_1$

$$b' = (1-r_2)\, e_1 - (1+r)\, e_0$$

and $S_{1.i} = \sum_0^i e_k$.

Again summing equation (6.2) over the values of i from 1 to i, we have

$$\sum_1^i (e_i - e_{i-1}) = r \sum_1^i S_{1.i} + r_1 \sum_{i=1}^i e_i + 2a \sum_{i=1}^i (i-1) + ib'$$

i.e. $e_i - e_0 = r \sum_1^i S_{1.i} + r_1 S_{1.i} - r_1 e_0 + 2a(0+1+2+\ldots+i-1) + ib'$

i.e. $e_i = r \sum_1^i S_{1.i} + r_1 S_{1.i} - e_0(r_1 - 1) + ai(i-1) + ib'$

or $e_i = rS_{2.i} + r_1 S_{1.i} + ai^2 + bi + C$ ... (6.3)

where $S_{2.i} = \sum_0^i S_{1.k}$

$C = e_0(1-r_2)$ and

$b = b' - a$.

The equation (6.3) is known as the internal regression of (i). This equation can be solved for r, r_1, a, b, and C by ordinary least squares method for fixed values of $S_{1.i}$ and $S_{2.i}$. This called a regression equation of e_i on $S_{1.i}$, $S_{2.i}$, i and

i^2. Let the least square estimates be denoted as $\hat{r}$, $\hat{r}_1$, $\hat{a}$, $\hat{b}$ and $\hat{C}$, then the relationship between the constants are

$$(\hat{d}_1,\hat{d}_2) = (1-\frac{\hat{r}_1}{2} \pm \sqrt{\hat{r} + \hat{r}_1^2/4})/(1-\hat{r}_1-\hat{r}),$$

$$\hat{A} = Z - 2\hat{a}/\hat{r}, \qquad \ldots \quad (6.4)$$

$$y_0 = \hat{A} + B_1 + B_2 = Z + C/(1-\hat{r}_1-\hat{r}) \quad \text{and}$$

$$y_1 = \hat{A} + B_1\hat{d}_1 + B_2\hat{d}_2$$

$$= Z + \hat{b}+\hat{a}+(\hat{r} + 1)\ \hat{C}/(1-\hat{r}_1 - \hat{r}) \ / \ (1-\hat{r}_1 - \hat{r})$$

The first two equations in (6.4) gives the values of $\hat{d}_1,\hat{d}_2$ and $\hat{A}$. Knowing the values of $\hat{d}_1,\hat{d}_2$ and $\hat{A}$, we can calculate the values of $\hat{B}_1$ and $\hat{B}_2$ by solving the last two equations. Thus all the constants of the curve (i) can be obtained.

Case 2:- Let x : -m, -(m-1), ... -1,0,1,... m. Then we have the definitions of $S_{1.i}$ and $S_{2.i}$ as

$$S_{1.i} = \sum_{0}^{i} e_k \quad \text{for} \quad i \geqslant 0,$$

$$S_{1.-1} = 0$$

$$S_{1.i} = -\sum_{-1}^{i+1} e_k \quad \text{for} \quad i \leqslant -2, \quad \text{and}$$

$$S_{2.i} = \sum_{0}^{i} S_{1.k} \quad \text{for} \quad i \geqslant 0,$$

$$S_{2.-1} = S_{2.-2} = 0, \qquad \ldots \qquad (6.5)$$

$$S_{2.i} = -\sum_{-i}^{i+1} S_{1.i} \quad \text{for} \quad i \leq -3 .$$

The internal regression and the relationship between the constants are the same as equations (6.3) and (6.4) respectively. Moreover we have

$$y_1 = A + B_1 d_1^{-1} + B_2 d_2^{-1} = Z + a - b + C.$$

Case 3:- Let x: $-m + \frac{1}{2}$, $-(m-1) + \frac{1}{2}$, ..., $-\frac{1}{2}$, $\frac{1}{2}$, ... $m - \frac{1}{2}$.

Then we have the following definitions of $S_{1.x}$ and $S_{2.x}$ as

$$S_{1.x} = \sum_{1/2}^{x} e_k \quad \text{for} \quad x \geq \tfrac{1}{2}$$

$$S_{1.-\frac{1}{2}} = 0$$

$$S_{1.x} = -\sum_{-1/2}^{x+1} e_k \quad \text{for} \quad x \leq -3/2 \quad \text{and} \ \ldots \qquad (6.6)$$

$$S_{2.x} = \sum_{1/2}^{x} S_{1.k} \quad \text{for} \quad x \geq \tfrac{1}{2}$$

$$S_{2.-\frac{1}{2}} = S_{2.-3/2} = 0,$$

$$S_{2.x} = -\sum_{-1/2}^{x+2} S_{1.k} \quad \text{for} \quad x \leq -5/2.$$

(Here the summation is from $\frac{1}{2}$, $\frac{1}{2} + 1$, $2 + \frac{1}{2}$, ... x or $-\frac{1}{2}$, $-1 -\frac{1}{2}$, ..., x + 1).

The internal regression is the same as equation (6.3) and the relationship between the constants are given as (d_1, d_2) and A same as equation (6.4), but

$$y_{\frac{1}{2}} = A + B_1 d_1^{\frac{1}{2}} + B_2 d_2^{\frac{1}{2}} = Z + (C + b/2 + a/4) / (1 - r - r_1) \quad \text{and}$$

$$y_{-\frac{1}{2}} = A + B_1 d_1^{-\frac{1}{2}} + B_2 d_2^{-\frac{1}{2}} = Z + C - b/2 + a/4 \ .$$

2.2 Consequences of regression equation (6.3).

(c.1): When $d_2 = 0$ in equation (i), the difference equation (iii) reduces to

$$y_{i+2} - d_1 \, y_{i+1} = A\,(1 - d_2).$$

Hence the internal regression of the type (6.3) reduces to

$$e_i = r_1 S_{1.i} + bi + C \qquad \ldots \qquad (6.7)$$

r_1, b and C can be calculated by well known normal equations obtained by considering $S_{1.i}$ as fixed values. The constants A, B_1 and d_1 can be obtained by the following relationship

$$\begin{aligned} \hat{d}_1 &= 1/(1 - \hat{r}_1) \\ \hat{A} &= Z - \hat{b}/\hat{r}_1 \quad \text{and} \quad \ldots \qquad (6.8) \\ \hat{B}_1 &= \hat{C}\,\hat{d}_1 + \hat{b}/\hat{r}_1 \end{aligned}$$

where $\hat{r}_1$, $\hat{b}$ and $\hat{C}$ are the least squares solution of equation

(6.7), when x : 0,1,2,..., n or x : -m,-(m-1),...0,1...m .
When x : $-m+\frac{1}{2}$, ..., $-\frac{1}{2}$, $\frac{1}{2}$,..., $m-\frac{1}{2}$, the constants are

$$\hat{d}_1 = 1/(1-\hat{r}_1)$$
$$A = Z-\hat{b}/\hat{r}_1 \quad \text{and} \quad \ldots \qquad (6.8')$$
$$e_{-\frac{1}{2}} = \hat{C} - \hat{b}/2 \quad .$$

(c.2) : When $d_1 = d_2$ in difference equation (iii), then the solution of (iii) is $y_x = A + (B_1 + B_2x)\, d^x$ and the internal regression will be the same as (6.3), but if in practice we have $r = -r_1^2/4$, then the above mentioned can be fitted and the relationship between the constants are same as in section (2.1).

(c.3) : When in practice, we find $r < -r_1^2/4$, then d_1 and d_2 are complex conjugate numbers and we find from the relation of the constants that B_1 and B_2 are also complex conjugate numbers. In this case, the regression equation is the same as (6.3), but the curve is written as

$$y_x = A + \delta^x \ (p \cos \mu x + q \sin \mu x) \quad \text{and}$$

the relations between the constants are

$$\delta^2 = w^2 + v^2 = 1/(1-r-r_1) \ ,$$

$$\tan\mu = v/w \quad \text{or} \quad \delta\cos\mu = w, \quad \delta\sin\mu = v ,$$

$$(d_1, d_2) = w \pm jv$$

$$= (1-r_1/2 \pm j\sqrt{-r-r_1^2/4})/(1-r-r_1), j=\sqrt{-1},$$

$$A = Z - 2a/r .$$

When x's are measured like 0,1,2,...,n or -m,...0,...m, then

$p=C\delta^2+2a/r$ and $q=\delta^2\left\{(r_1/2 + r)\, C\delta^2 + b+(r-r_1)\, a/r\right\}/v.$

When x's are measured like $-m + \frac{1}{2}$, ..., $-\frac{1}{2},\frac{1}{2}$,...$m-\frac{1}{2}$, then

$$p = \frac{\left\{(a+4C)\, r\delta-8a\right\}(\delta+1) + 2rb\delta(\delta-1)}{8r\, \delta^{\frac{1}{2}}\, \cos\mu/2} \quad \text{and}$$

$$q = \frac{\left\{(a+4C)\, r\delta-8a\right\}(\delta-1) + 2rb\,\delta(\delta+1)}{8r\, \delta^{\frac{1}{2}}\, \sin\mu/2} .$$

2.3 : To fit $y = A + Hx + B_1d_1^x + B_2d_2^x$.

The difference equation (iv) can be rewritten as

$$(e_{i+2}- e_{i+1})-(e_{i+1}- e_i) = r_2e_{i+2}- r_1e_{i+1} + 6hi + 2a' \quad \ldots (6.9)$$

where

$$H = -6\hat{h}/ (\hat{r}_2 - \hat{r}_1)$$

$$A = Z - (2\hat{a} + \hat{H}\hat{r}_2) / (\hat{r}_2 - \hat{r}_1)$$

$$\hat{a}' = a + 3\hat{h},$$

$$d_1 + d_2 = (2-\hat{r}_1)/(1-\hat{r}_2) \ ,$$

$$d_1 d_2 = 1/(1-\hat{r}_2) \quad \text{and}$$

$$y_i - Z = e_i, \qquad Z \text{ being any value.}$$

Then summing (6.9) over the values of i from 0 to i-2, we have

$$e_i - e_{i-1} = rS_{1.i} + r_1 e_i + 3h(i-2)\ (i-1) + (i-1)\ 2a' + b' \ ,$$

where r and b' are defined in (2.1).

Again summing the above equation over the values of i from 1 to i, we have, the regression equation of e_i on $S_{2.i}$, $S_{1.i}$, i^3, i^2, and i as

$$e_i = rS_{2.i} + r_1 S_{1.i} + hi^3 + ai^2 + bi + C \quad \ldots \qquad (6.10)$$

where r, $S_{1.i}$, $S_{2.i}$ are the same as defined in (2.1) and the relationship between the constants are (when x:0,1,2,...n)

$$H = -6\hat{h}/\hat{r},$$

$$H + A = Z - (2\hat{a} + H\hat{r}_1)/\hat{r}, \quad \ldots. \qquad (6.11)$$

$$C = e_0\ (1 - \hat{r}_2) \quad \text{and}$$

$$\hat{b} + \hat{a} + \hat{h} = (1-\hat{r}_2)\ e_1 - (1 + \hat{r})\ e_0$$

When x : $-m, -(m-1), \ldots 0, \ldots, m$,

$$H = -6\hat{h} / \hat{r},$$

$$H + A = Z - (2\hat{a} + H\hat{r}_1)/\hat{r}, \quad \ldots \quad (6.12)$$

$$C = e_0(1-\hat{r}_2) \quad \text{and}$$

$$e_{-1} = -\hat{h} + \hat{a} - \hat{b} + \hat{C}.$$

When x : $-m + \frac{1}{2}, \ldots -\frac{1}{2}, \frac{1}{2}, \ldots m-\frac{1}{2}$,

$$H = -6\hat{h} / \hat{r}$$

$$H + A = Z - (2\hat{a} + H\hat{r}_1)/\hat{r},$$

$$e_{\frac{1}{2}} = (\hat{h}/8 + \hat{a}/4 + \hat{b}/2 + \hat{C}) / (1-\hat{r}-\hat{r}_1) \quad \ldots \quad (6.13)$$

and $e_{-\frac{1}{2}} = (-\hat{h}/8 + \hat{a}/4 - \hat{b}/2 + \hat{C})$.

Note that the constants $\hat{a}, \hat{b}, \hat{c}, \hat{h}, \hat{r}$ and $\hat{r}_1$ are the solutions of the normal equations obtained by considering $S_{1.i}$, $S_{2.i}$ as fixed values.

2.4 : Consequences of the regression equation (6.10):-

(C.4):- When $d_2=0$ in equation (ii), the difference equation (iv) reduces to $y_{i+2} - d_1 y_{i+1} = A(1-d_1)+Hi(1-d_1)+H(2-d_1)$.

Hence the internal regression and the relationship between the constants are respectively as

$$e_i = r_1 S_{1.i} + hi^2 + bi + C \quad \ldots \quad (6.14)$$

and $d_1 = 1/(1-\hat{r}_1)$, $H = -2\hat{h}/\hat{r}_1$, $A = Z + (\hat{h}+H)/\hat{r}_1 - \hat{b}/\hat{r}_1 \ldots (6.15)$

$e_0(1-r_1) = \hat{C}$, when $x = 0,1,2,\ldots n$ or $x = -m,\ldots,0,\ldots m$;

when $x = -m + \frac{1}{2},\ldots -\frac{1}{2},\frac{1}{2},\ldots m-\frac{1}{2}$,

$e_{-\frac{1}{2}} = \hat{h}/4 - \hat{b}/2 + \hat{C}$ or $e_{\frac{1}{2}}\ (1-\hat{r}_1) = \hat{h}/4 + \hat{b}/2 + \hat{C}$,

$d_1 = 1/(1-\hat{r}_1)$, $H = -2\hat{h}/\hat{r}_1$, $A = Z + (\hat{h} + H - \hat{b})/\hat{r}_1 \ldots (6.15')$

(C.5). When $d_1 = d_2$ in the difference equation (iv), then the solution of (iv) is $y_x = A + Hx + (B_1 + B_2 x)d^x$ and the internal regression will be the same as (6.10), but if in practice, we have $r = -r_1^2/4$, then the above mentioned curve can be fitted and the relationship between the constants is the same as (2.2).

(C.6). When $r + r_1^2/4 < 0$, then the same argument applies as (C.3). The curve (ii) is then broken up as

$y_x = A + Hx + \delta^x (p \cos \mu x + q \sin \mu x)$ and the relation-ship between the constants can easily be derived from (2.3).

3. Numerical examples:-

3.1. Here we shall first fit the logistic curve which was fitted by H.O. Hartley [10] to verify our new type of formula derived in (C.1).

x	y	w=1000/y	$e_i = w_i - Z$	$S_{1.i}$	Expected
-5.5	5.308	1884	1277.91667	-1182.58335	1864.44
-4.4	7.240	1381	774.91667	-407.66668	1382.94
-3.5	9.368	1038	431.91667	-24.24999	1026.78
-2.5	12.866	777	170.91667	195 .16666	765.49
-1.5	17.069	586	-20.08333	175 .08333	573.81
-0.5	23.192	431	-175.08333	0.00000	433.19
0.5	31.443	318	-288.08333	-288.08333	330.03
1.5	38.558	259	-347.08333	-635.16666	254.35
2.5	50.156	199	-407.08333	-1042.24999	198.83
3.5	62.948	159	-447.08333	-1489.33332	158.10
4.5	75.995	132	-474.08333	-1963.41665	128.22
5.5	91.972	109	-497.08333	-2460.49998	106.30
			+0.00004	-9074.49998	

Here $Z = 606.08333$ and $\bar{S}_{1.i} = -756.20833$.

Corrected sum of squares and products (S.S.P.).

	$S_{1.i}$	x	e
$S_{1.i}$	8471707.84965	-23779.91619	1760647.32966
x		143	-20451.5

Then the internal regression is

$$e_i = (-0.363123)\, S_{1.i} + (-203.40208)\, i + (-274.596637)$$

Hence by the relationship of the constants, the curve is $w = 45.936827 + 331.684567(0.7336095)^x$, while Hartley has obtained the curve as $w = 50.4 + 323.4\ (0.7336)^x$.

The difference between the constants A and B will be due to the different ways of obtaining A and B.

3.2. We have fitted the curve (1) to the data of Makeham taken from "Life contingencies" [35]

age	$\mu_x = y_x$	w_x	x	$e_x = y_x - Z$	$S_{1 \cdot x}$	$S_{2 \cdot x}$	Expected
3	0.017870	0.01337357	-12	0.0113452	0.0001728	-0.1107840	0.0173174
4	0.01379	0.01699213	-11	0.0072652	0.0074380	-0.1033460	0.0139201
5	0.01142	0.02021090	-10	0.0048952	0.0123332	-0.0910128	0.0115536
6	0.00925	0.02464266	-9	0.0027252	0.0150584	-0.0759544	0.0093249
7	0.00748	0.03016682	-8	0.0009552	0.0160136	-0.0599408	0.0075084
8	0.00007	0.03687227	-7	-0.0004548	0.0155588	-0.0443820	0.0060752
9	0.00502	0.04429194	-6	-0.0015048	0.0140540	-0.0303280	0.0049903
10	0.00428	0.05167156	-5	-0.0022448	0.0118092	-0.0185188	0.0042153
11	0.00388	0.05674309	-4	-0.0026448	0.0091644	-0.0093544	0.0037098
12	0.00359	0.06108189	-3	-0.0029348	0.0062296	-0.0031248	0.0034337
13	0.02342	0.06388367	-2	-0.0031048	0.0031248	0	0.0033480
14	0.00340	0.06403909	-1	-0.0031248	0	0	0.0034159
15	0.00353	0.06147777	0	-0.0029948	-0.0029948	-0.0029948	0.0036031
16	0.00378	0.05721571	1	-0.0027448	-0.0057396	-0.0087344	0.0038792
17	0.00414	0.05205420	2	-0.0023848	-0.0081244	-0.0168588	0.0042177
18	0.00458	0.04686891	3	-0.0019448	-0.0100692	-0.0269280	0.0045934
19	0.00504	0.04240664	4	-0.0014848	-0.0115540	-0.0384820	0.0049884
20	0.00550	0.03867337	5	-0.0010248	-0.0125788	-0.0510608	0.0053858
21	0.00592	0.03573980	6	-0.0006048	-0.0131836	-0.0642444	0.0057728
22	0.00629	0.03344490	7	-0.0002348	-0.0134184	-0.0776628	0.0061395
23	0.00659	0.03172719	8	0.0000652	-0.0133532	-0.0910160	0.0064786
24	0.00682	0.03045915	9	0.0002952	-0.0130580	-0.1040740	0.0067853
25	0.00701	0.02943458	10	0.0004852	-0.0125728	-0.1166468	0.0070568
26	0.00716	0.02861848	11	0.0006352	-0.0119376	-0.1285844	0.0072691
27	0.00729	0.02790974	12	0.0007652	-0.0111724	-0.1397568	0.0074908
	0.16312	1.0000001	0	0.0000000			

Here $Z = 0.0065248$.

$\sum we = -0.00108075$, $\sum wS_{1.i} = -0.00099$, $\sum wS_{2.i} = -0.04181055$,

$\sum wx = 0.18862962$, $\sum wx^2 = 36.58095808$.

weighted corrected S.S.P.

	x	x^2	S_1	S_2	e
x	36.54537695	41.90536574	-0.05647117	-0.10596896	-0.00258384
x^2		1662.28565062	-0.03173886	-1.60297659	0.08529805
S_1			0.00010670	0.00014647	0.00000110
S_2				0.00173022	-0.00007890

Hence the internal regression is (by applying Doolittle technique)

$$e_x=(-0.0393712)S_{2.x}+(-0.22281618)S_{1.x} + 0.0000231\ x^2$$

$$+(-0.00055565)x +(-0.00368768).$$

From this, we have $(d_1, d_2) = 0.88054128 \pm j\ 0.130086295 = w + jv$.

By the use of the relationship between the constants (C.3), we have the curve (i) as

$$y_x=0.00769825 + (0.89009852)^x \left\{(-0.00409511) \cos\mu x + (-0.00163807) \sin\mu x\right\} \text{ where } \mu = 8^\circ\ 24.2'.$$