Journal of Applied Statistics, Vol. 14, No. 1, 1987

A non-normal class of distribution function for dose-binary response curve*

M. SINGH, *International Crops Research Institute for the Semi-Arid Tropics, Patancheru 502 324, Andhra Pradesh*

SUMMARY *A non-normal class of distribution (Edgeworth Series distribution) function in three and four parameters has been considered for dose-binary response relationship. This class accounts for the non-normality (expressed in terms of skewness and kurtosis) present in the relationship in addition to the usual location and scale parameters (generally considered by two parameter models). We present the maximum likelihood method of estimation of the parameters and test of probit (normal distribution) hypothesis. Edgeworth Series distribution when fitted to the data of Milicer & Szczotka (1966) showed an excellent closeness to the observed values, significant improvement over probit and logit fit (Aranda-Ordaz, 1981), and better fit compared to Prentice (1976) model.*

1 Introduction

The binary response data reflecting the effect of an explanatory variable (dose) are often analysed using the relation given by a distribution function. If d is the dose applied on an experimental unit, the probability of response can be represented by

$$\begin{aligned} P(d) &= \int_{-\infty}^{(d-\theta_1)/\theta_2} f(t, \theta_3, \theta_4, \ldots)dt \\ &= F(d, \theta_1, \theta_2, \theta_3, \theta_4, \ldots) \end{aligned} \tag{1}$$

where functions $f(.)$ and $F(.)$ are probability density and corresponding distribution functions respectively. The parameters θ_1 (for location), θ_2 (for scale), $\theta_3, \theta_4, \ldots$ etc. require estimation. Two particular forms of $F(.)$ representing normal probability density function $F(d, \mu, \sigma) = \Phi\ (Z)$ known as probit transformation and logistic function $F(d, \mu, \sigma) = \exp(Z)/(1+\exp\ (Z))$, $Z = (d-\mu)/\sigma$, known as logit transforma

* Submitted as J.A. No. 628 by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.

tion have been extensively used in explaining real data (see, Bliss, 1935; Finney, 1971; Cox, 1970, etc.). Prentice (1976) considered a class of distribution functions based on four parameters representing shapes along with location and scale. Two families of transformations were discussed by Aranda-Ordaz (1981). Morgan (1985) extended the logit model by including cubic and quartic terms in dose variate and applied on several data sets. Here we considered the Edgeworth Series distribution function (Kendall & Stuart, 1977, pp. 168–169; Subrahmaniam, 1969) which can show departure from probit transformation (resulting into a non-normal distribution function) and is not a special case or extension of Prentice (1976) and Morgan (1985). The Edgeworth Series Distribution (ESD) model can be used to adjust for departure (in shapes reflected in skewness and kurtosis) from normality. The model, maximum likelihood estimates of its parameters and also the estimate of effective dose (ED_γ) required for a given percentage (γ) of response are in Section 2. When applied to the data of Milicer & Szczotka (1966) in Section 3 ESD model provided remarkably closer fit compared to that discussed by Aranda-Ordaz (1981) on probit and logit models and also by Prentice (1976).

2 Edgeworth series distribution (ESD) model

2.1 *Model*

We confine to the situation of only four parameters. This non-normal distribution function with parameters (μ, σ, λ_3, λ_4) is given by

$$P(d)=\int_{-\infty}^{z} f(t, \lambda_3, \lambda_4)dt$$

where

$$f(t, \lambda_3, \lambda_4)=\phi(t)\{1+\lambda_3 H_3(t)/6+\lambda_4 H_4(t)/24+\lambda_3^2 H_6(t)/72\}$$

Also

$$P(d)=\Phi(z)-(\lambda_3 H_2(z)/6+\lambda_4 H_3(z)/24+\lambda_3^2 H_5(z)/72)\Phi(z)$$

$$z=(d-\mu)/\sigma,\ \phi(t)=(1/(2\pi)^{\frac{1}{2}})e^{-t^2/2};\ \Phi(z)=\int_{-\infty}^{z}\phi(t)dt$$

$H_i(t)(i=2, 3,\ldots, 6)$ is Hermite Polynomial of order i in t.

$$H_2(t)=t^2-1,\ H_3(t)=t^3-3t,\ H_4(t)=t^4-6t^2+3,\ H_5(t)=t^5-10t^3+15t$$
$$H_6(t)=t^6-15t^4+45t^2-15$$

$H_i(t)$ can be defined in general, by

$$(-1)^i d^i\ \phi(t)/dt^i=H_i(t)\phi(t)$$

In above μ, σ, λ_3 and λ_4 are parameters of location, scale, skewness and kurtosis respectively. If $\lambda_3=\lambda_4=0$, then we get normal probability distribution (probit). Thus ESD model allows the adjustment for departure from normality by accounting for shape parameters λ_3 and λ_4.

2.2 *Estimation of parameters*

Let m represent the number of individuals (units) out of n responded at dose d_r ($r=1\ldots k$), k is number of distinct doses considered in the experiment. The log-likelihood for

$$\theta=(\theta_1, \theta_2, \theta_3, \theta_4)$$

where

$$(\theta_1=\mu,\ \theta_2=\sigma,\ \theta_3=\lambda_3,\ \theta_4=\lambda_4$$

for convenience of notation), is except for a constant term (independent of θ)

$$l=\sum_{r=1}^{k}(m_r\log(P(d_r))+(n_r-m_r)\log(Q(d_r))$$

where $Q(d_r)=1-P(d_r)$

The maximum likelihood estimates (m.l.e.) of θ can be obtained by solving the following equations iteratively (see Rao, 1973, p. 366).

$$\partial l/\partial\theta_i=\sum_{i=1}^{k}(m_r/P(d_r)-(n_r-m_r)/Q(d_r))\partial P(d_r)/\partial\theta_i=0$$

for $i=1, 2, 3, 4$.

We note

$$\partial P(d_r)/\partial\mu=(-1/\sigma)f(Z_r),\ Z_r=(d_r-\mu)/\sigma$$
$$\partial P(d_r)/\partial\sigma=(-Z_r/\sigma)f(Z_r)$$
$$\partial P(d_r)/\partial\lambda_3=-(H_2(Z_r)/6+H_5(Z_r)/36)\phi(Z_r)$$
$$\partial P(d_r)/\partial\lambda_4=-H_3(Z_r)\phi(Z_r)/24$$

The information matrix for θ denoted by l_θ has (i, j)th element

$$\sum_{r=1}^{k}(n_r/(P(d_r)Q(d_r)))(\partial P(d_r)/\partial\theta_i)(\partial P(d_r)/\partial\theta_j)$$

for $i, j=1, 2, 3, 4$

The iterative procedure to get estimate of θ can also be seen from Prentice (1976) using expressions for $\partial l/\partial\theta_i$, l_θ and an initial solution.

2.3 Estimation of effective dose ED (γ)

For given values of θ_1, θ_2, θ_3, θ_4 (as their m.l.e.) $ED(\gamma)$ for a given percentage γ of response can be estimated by $\hat{x}$, a solution of

$$S(x,\hat{\theta})=\Phi(Z)-(\hat{\theta}_3H_2(Z)/6+\hat{\theta}_4H_3(Z)/24+\hat{\theta}_3^2H_5(Z)/72)\phi(Z)=\gamma/100$$
$$Z=(x-\hat{\theta}_1)/\hat{\theta}_2$$

We used Newton-Ralphson method to get the solution $Z^\star$ for above equation in Z, choosing the initial value of Z as $\Phi^{-1}(\gamma/100)$ and hence $\hat{x}=\hat{\theta}_1+\hat{\theta}_2Z^\star$. $\hat{\theta}_i$ is the m.l.e. of θ_i $(i=1, 2, 3, 4)$.

The asymptotic variance Va($\hat{x}$) of $\hat{x}$ can be written as

$$\mathrm{Va}(x)=\sum_i\sum_j(\partial x/\partial\theta_i)(\partial x/\partial\theta_j)\ \mathrm{Cov}(\hat{\theta}_i,\hat{\theta}_j)$$

where the solution $\hat{x}$ of $S(x,\hat{\theta})=\gamma/100$ can be written as function $x(.)$

$$\hat{x}=x(\gamma,\hat{\theta})$$

in order to evaluate $\partial x/\partial\theta_i$.

Now partial differentiation of $S(x,\theta)=\gamma/100$ with respect to x and θ_i gives

$$(\partial S/\partial x)(\partial x/\partial\theta_i)+(\partial S/\partial\theta_i)=0$$

or

$$\partial x/\partial\theta_i=-(\partial S/\partial\theta_i)/(\partial S/\partial x)$$

We know

$$\partial S/\partial \theta_i$$

and

$$\partial S/\partial x = (\partial/\partial x) \int_{-\infty}^{x} f(t)dt = (1/\sigma)f(x)$$

Thus

$$\partial x/\partial \theta_i = -(\sigma/f(x))\ \partial S/\partial \theta_i$$

Therefore

$$\mathrm{Va}(\hat{x}) = (\sigma^2/f^2(x)) \sum_i \sum_j (\partial S/\partial \theta_i)(\partial S/\partial \theta_j)\ \mathrm{Cov}(\hat{\theta}_i, \hat{\theta}_j)$$

and a large sample confidence bound for x can be obtained using asymptotic normal likelihood estimates $\hat{\theta}$'s.

3 Illustration

We fitted the three and four parameter ESD model of Section 2 using the data of Milicer & Szczotka (1966) on the records of 3918 Warsaw girls in 1963 giving individuals age and whether or not she had reached menarche. Table 1 (reproduced from Aranda-Ordaz 1981) gives the mean age of groups (years), number of girls, number menstruated and the comparison with probit, logit and Prentice (1976) models in terms of fitted values of proportions, maximum likelihood function $l(l_1, l_2, l_3, l_4, l_5)$ and likelihood ratio chi-square values (Rao, 1973, p. 414; Kolakowski & Bock, 1981).

The maximum likelihood function values for ESD models $l_1 = -813.29$ adjusted for parameter of skewness, $l_2 = -813.10$ adjusted for parameters of both skewness and kurtosis in addition to location and scale; for probit $l_3 = -817.74$, for logit (Aranda-Ordaz took $\lambda = 0.325$ in his transformation) $l_4 = -817.58$, and for Prentice, $l_5 = -815.77$. Thus comparing $2(l_1 - l_3) = 8.9$ with chi-square table value on 1 d.f. shows more significant improvement over probit and logit models compared to Prentice (1976) model with three parameters (Table 1).

The maximum likelihood estimation of the parameters in ESD model with only three parameters are $\hat{\mu} = 13.008$, $\hat{\sigma} = 1.1022$, $\hat{\lambda}_3 = 0.3989$ with variance-covariance matrix estimate as

$$\begin{array}{cc} & \begin{array}{ccc} \hat{\mu} & \hat{\sigma} & \hat{\lambda}_3 \end{array} \\ \begin{array}{c} \hat{\mu} \\ \hat{\sigma} \\ \hat{\lambda}_3 \end{array} & 10^{-3} \begin{bmatrix} 1.4832 & & \\ 0.1705 & 1.2066 & \\ -0.2427 & 0.5993 & 1.7734 \end{bmatrix} \end{array}$$

The maximum likelihood estimates for four parameters ESD model are $\hat{\mu} = 13.007$, $\hat{\sigma} = 1.1038$, $\hat{\lambda}_3 = 0.4135$, $\hat{\lambda}_4 = 0.1987$ with variance and covariance matrix

$$\begin{array}{cc} & \begin{array}{cccc} \hat{\mu} & \hat{\sigma} & \hat{\lambda}_3 & \hat{\lambda}_4 \end{array} \\ \begin{array}{c} \hat{\mu} \\ \hat{\sigma} \\ \hat{\lambda}_3 \\ \hat{\lambda}_4 \end{array} & 10^{-3} \begin{bmatrix} 1.4985 & & & \\ 0.1538 & 1.3546 & & \\ -0.4712 & 1.2681 & 9.8465 & \\ -0.7491 & 1.9481 & 25.6970 & 81.6430 \end{bmatrix} \end{array}$$

The values of likelihood ratio chi-square for goodness of fit on 22 degrees of freedom is 13.98 for three parameters and 13.60 for four parameter case showing an excellent fit to the data. However, the four parameters fit does not show significant improvement over three parameters in this case, but may be of importance in those distributions arising in practice with non-zero coefficient of kurtosis.

TABLE 1. Age of menarche (in years) in 3918 Warsaw girls (Milicer & Szczotka, 1966) and fitted values from ESD model with three and four parameters, Probit, logit and Prentice Models

Case	Age	N	M	OBS	ESD3	ESD4	Probit	Logit	P3PD
1	9.21	376	0	0.0000	0.0001	0.0001	0.0003	0.0020	0.0002
2	10.21	200	0	0.0000	0.0001	0.0001	0.0054	0.0103	0.0032
3	10.58	93	0	0.0000	0.0032	0.0039	0.0134	0.0187	0.0082
4	10.83	120	2	0.0167	0.0112	0.0114	0.0234	0.0278	0.0150
5	11.08	90	2	0.0222	0.0263	0.0260	0.0392	0.0413	0.0270
6	11.33	88	5	0.0568	0.0519	0.0507	0.0626	0.0609	0.0469
7	11.58	105	10	0.0952	0.0904	0.0882	0.0957	0.0889	0.0784
8	11.83	111	17	0.1532	0.1430	0.1401	0.1402	0.1278	0.1250
9	12.08	100	16	0.1600	0.2095	0.2065	0.1970	0.1806	0.1889
10	12.33	93	29	0.3118	0.2876	0.2852	0.2658	0.2489	0.2699
11	12.58	100	39	0.3900	0.3736	0.3724	0.3451	0.3226	0.3637
12	12.83	108	51	0.4722	0.4630	0.4634	0.4319	0.4284	0.4639
13	13.08	99	47	0.4747	0.5512	0.5534	0.5221	0.5299	0.5631
14	13.33	106	67	0.6321	0.6344	0.6382	0.6111	0.6290	0.6534
15	13.58	105	81	0.7714	0.7100	0.7146	0.6948	0.7182	0.7321
16	13.83	117	88	0.7521	0.7762	0.7807	0.7692	0.7931	0.7970
17	14.08	98	79	0.8061	0.8321	0.8359	0.8322	0.8521	0.8487
18	14.33	97	90	0.9278	0.8778	0.8804	0.8830	0.8966	0.8886
19	14.58	120	113	0.9417	0.9138	0.9149	0.9218	0.9288	0.9187
20	14.83	102	95	0.9314	0.9411	0.9409	0.9499	0.9515	0.9411
21	15.08	122	117	0.9590	0.9607	0.9598	0.9693	0.9672	0.9575
22	15.33	111	107	0.9640	0.9746	0.9731	0.9821	0.9779	0.9695
23	15.58	94	92	0.9787	0.9838	0.9822	0.9900	0.9852	0.9781
24	15.83	114	112	0.9825	0.9899	0.9885	0.9946	0.9901	0.9844
25	17.58	1049	1049	1.0000	0.9998	0.9998	1.0000	0.9994	0.9985
Likelihood ratio					13.98	13.60	22.92	27.25	18.98
d.f.					22	21	23	23	22
Log likelihood value					−813.29	−813.10	−817.84	−819.92	−815.77

OBS stands for observed proportion; ESD3, P3PD for fitted proportion using ESD model and Prentice (1976) model with three parameters, ESD4 for ESD model with four parameters.

The estimates of effective age values for given percentage (γ) of response along with their asymptotic 95% confidence limits are given below in Table 2.

TABLE 2. Effective age (years) and 95% asymptotic confidence limits

	Three parameter case			Four parameter case		
	Effective	Asymptotic limit		Effective	Asymptotic limit	
γ	Age (years)	Lower	Upper	Age (years)	Lower	Upper
50%	12.933	12.853	13.012	12.930	12.843	13.018
95%	14.934	14.779	15.089	14.941	14.768	15.113
99%	15.837	15.620	16.054	15.908	15.533	16.282

Discussion

The procedure illustrated in the previous section shows how an Edgeworth Series distribution even with finite number of terms of the series can capture the departure

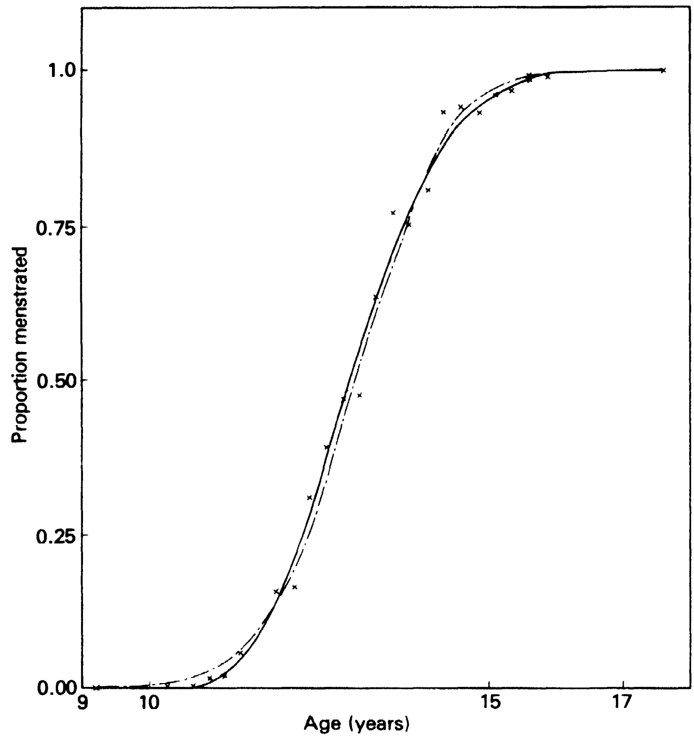

FIG. 1.

from probit upto a significant extent and outperform its several rivals. A graphical display is also given in Fig. 1 only for the value of observed and fitted proportions using three parameters ESD and Probit models for the sake of clarity.

However, the inclusion of finite number of terms of Hermite polynomial in the density may lead to negative values of density function of ESD model particularly when the values λ_3 and λ_4 parameters are large. But for moderate (departure from) improvement over Probit model the ESD density can be useful. This model may also provide non-unique solution to effective dose values but the Newton-Ralphson method to find out real root yielded an acceptable solution when applied to the data illustrated.

It may also be noticed that the ESD model can be used to provide a test for skewness ($\lambda_3=0$) and also for kurtosis ($\lambda_4=0$) of the distribution under investigation. The present case leads to the rejection of the hypothesis $\lambda_3=0$ but not of $\lambda_4=0$.

Furthermore, with regard to the data considered in the paper one might argue for the over collection of the data on the tails, under the assumption or prior knowledge that no girl less than 10.5 years of age would have reached menarche and all girls older than 15.83 years would menstruate. Fitting ESD model to the reduced data ignoring the cases (1, 2, 3) below 10.6 years and case (25) above 17.5 years (Table 1), we get no remarkable change in the estimates of the location and scale parameters estimates ($\hat{\mu}=13.001$, $\hat{\sigma}=1.109$) and log likelihood value of -812.68 but approximately 12 percent change in skewness parameter ($\hat{\lambda}_3=0.352$). However, better assessment of the model will require the processing of some more data sets.

Acknowledgement

The author sincerely thanks several reviewers for their encouraging comments and to Mr C. P. Jaiswal for his secretarial help.

Correspondence: M. Singh, International Crops Research Institute for the Semi-Arid Tropics, Patancheru 502 324, Andhra Pradesh, India.

REFERENCES

ARANDA-ORDAZ, F.J. (1981) On two families of transformations to additivity for binary response data, *Biometrika*, 68, pp. 357–63.

BLISS, C.I. (1935) The calculation of dose-mortality curve, *Annals of Applied Biology*, 22, pp. 134–67.

COX, D.R. (1970) *The Analysis of Binary Data* (Methuen, London).

FINNEY, D.J. (1971) *Probit Analysis*, 3rd edn (Cambridge University Press).

KENDALL, M.G. & STUART, A. (1977) *The Advanced Theory of Statistics*, 4th edn, vol. 1 (London, Charles Griffin).

KOLAKOWSKI, D. & BOCK, R.D. (1981) A multivariate generalisation of probit analysis, *Biometrics*, 37, pp. 541–551.

MORGAN, B.J.T. (1985) The Cubic Logistic Model for Quantal Assay Data, *Applied Statistics*, 34, pp. 105–113.

MILICER, H. & SZCZOTKA, F. (1966) Age at menarche in Warsaw girls in 1965, *Human Biology*, 38, pp. 199–203.

PRENTICE, R.L. (1976) Generalisation of Probit and Logit Models, *Biometrics*, 32, pp. 761–768.

RAO, C.R. (1973) *Linear Statistical Inference with Applications* (New Delhi, Wiley Eastern).

SUBRAHMANIAM, K. (1969) Order Statistics from a class of non-normal distributions, *Biometrika*, 56, pp. 415–428.