THE DESIGN OF QUANTAL RESPONSE EXPERIMENTS AND THE MODELLING OF QUANTAL RESPONSE EXPERIMENTS OVER TIME

by

RANDY RUDOLF SITTER

B.Sc., The University of British Columbia, 1984

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Statistics Department

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August 1986

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Statistics

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date Aug 14/86

ABSTRACT

The problem of designing a quantal response experiment when estimation of the median effective dose (ED50) is of main interest is examined. The asymptotic variances of the maximum likelihood estimators of the ED50 for various 3 and 5 point designs, using the logit model, are compared to the minimum possible which is achieved with an inadvisable 1 point design (Chernoff[5]). Alternate criteria for choosing a design that attempt to incorporate goodness-of-fit of the model are then examined.

The modelling of quantal response experiments observed over time is also considered. A growth-curve approach to this problem was suggested by Carter and Hubert[3], and applied to a data set. The feasibility of this approach is discussed, and a simpler, more direct approach is proposed. The two models are applied to the presented data set, and the resulting fits are compared. The model proposed here appears to fit the data better. Inference about the ED50 using the two models is also compared.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

I thank Dr. A. J. Petkau for his guidance and encouragement in the producing of this thesis. I also thank Dr. Harry Joe for his suggestions and comments. This work was supported by the Natural Sciences and Engineering Research Council of Canada as well as by teaching and research assistanceships from the Department of Statistics, University of British Columbia.

I. Designs For Quantal Response Experiments Based On Minimum Variance

1. Description of Quantal Response Models

A typical quantal response problem consists of an observation y at a dose level x classified into two categories, response and non-response, with probabilities $\mathrm{P}(x)$ and $1-\mathrm{P}(x)$, respectively. Take n_i independent observations of this type at k dose levels observing the number of responses r_i, $i = 1, ..., k$. If we then assume $\mathrm{P}(x) = F(x \mid \underline{\theta})$ where $F(x \mid \underline{\theta})$ is a distribution function, and $\underline{\theta}$ is a vector of parameters, the log likelihood function is given by

$$L(\underline{\theta}) = \sum_{i=1}^{k} \left\{ r_i \log F(x_i \mid \underline{\theta}) + (n_i - r_i) \log[1 - F(x_i \mid \underline{\theta})] \right\}.$$

The maximum likelihood estimator (MLE) $\hat{\underline{\theta}}$ can then be found for $\underline{\theta}$. Depending on the choice of $F(x \mid \underline{\theta})$, the maximum likelihood equations, obtained by setting the derivative of $L(\underline{\theta})$ with respect to $\underline{\theta}$ to 0 and solving for $\underline{\theta}$, may not have an explicit solution. In this case the MLE must be found numerically using an iterative method such as Newton-Raphson.

The most commonly used models, for quantal response problems, are the probit model (Finney[7]) and logit model (Berkson[1]). The probit model is obtained by assuming

$$F(x \mid \underline{\theta}) = \Phi(\alpha + \beta x)$$

and the logit by assuming

$$F(x \mid \underline{\theta}) = \Psi(\alpha + \beta x),$$

where

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

and

$$\Psi(t) = [1 + e^{-t}]^{-1},$$

with $-\infty < t < +\infty$ and $-\infty < \alpha < +\infty$, $0 < \beta < +\infty$ the unknown parameters corresponding to $\underline{\theta}$.

Often the main interest in modelling quantal response curves is the estimation of the particular dose level at which the probability of response is 50%. Generally the dose level effecting ϑ% of the test subjects is denoted by EDϑ. Thus the 50% response dose is denoted by ED50. For "reasonable designs" (design points not too distant from the ED50) estimation of the ED50 using the probit and logit models produces similar results. Chapter I and Chapter II will deal with the logit model, but a similar development could be carried out for the probit model.

For estimation of the ED50 using the logit model a simple reparameterization is convenient. Letting $\mu = -\frac{\alpha}{\beta}$ and $\sigma = \frac{1}{\beta}$, yields $\Psi(\alpha+\beta x) = \Psi(\frac{x-\mu}{\sigma})$, and μ is the ED50. The information matrix for the MLE, $\underline{\hat{\theta}} = (\hat{\mu}, \hat{\sigma})^T$ is

$$I(\underline{\theta}) = \frac{n}{\sigma^2} \begin{pmatrix} \sum_{i=1}^{k} \lambda_i \psi(z_i) & \sum_{i=1}^{k} \lambda_i z_i \psi(z_i) \\ \sum_{i=1}^{k} \lambda_i z_i \psi(z_i) & \sum_{i=1}^{k} \lambda_i z_i^2 \psi(z_i) \end{pmatrix} \qquad (1.1)$$

where $z_i = (x_i - \mu)/\sigma$, $n = \sum_{i=1}^{k} n_i$, $\lambda_i = n_i/n$, and $\psi(t) = e^t/(1+e^t)^2$, and thus the asymptotic variance of $\hat{\mu}$ is given by

$$V(\hat{\mu}) = I_{11}^{-1}(\underline{\theta}).$$

A question which immediately comes to mind when considering such a quantal response experiment is the optimal experimental design in terms of what dose levels to use, the number of dose levels to use, and how to allocate subjects to these dose levels.

2. Optimal Design for Estimation of ED50

In considering this problem it is useful to first look at a linear regression model with

$$Y = \beta_1 x_1 + ... + \beta_k x_k + e,$$

where e is a random variable, with mean 0 and variance v^2, and for different observations Y_i the corresponding e_i's are independent. One formulation of the optimal design problem is the following: for $\underline{x} = (x_1, ..., x_k)$ in a specified set S, select n points of S, $(\underline{x}_1, ..., \underline{x}_n)$, so as to yield a minimum variance unbiased estimator of

$$\phi = \sum_{i=1}^{k} a_i \beta_i = \underline{a}^T \underline{\beta},$$

where $a_1, ..., a_k$ are known constants. Elfving[6] gave the following graphical solution to this problem (see also Chernoff[5]): Let S^* be the convex set generated by the points of S and S^-, the reflection of S about the origin.

Then $\underline{z}$, the point at which the ray from the origin through $\underline{a}$ intersects S^*, represents the solution in the following way. If $\underline{z}$ is a convex combination of points $\underline{x}_i$ of S or $-\underline{x}_i$ of S^- with weights λ_i, assign $n\lambda_i$ observations to experimental level $\underline{x}_i$. The variance of the resulting estimate of ϕ will then be

$$\frac{v^2}{n}\frac{\|\underline{a}\|^2}{\|\underline{z}\|^2}.$$

If the $n\lambda_i$ are not integers, this is not an exact solution and slightly understates the achievable variance.

Chernoff[5] uses this method to solve a more general problem, which includes the optimal design problem of interest here. In the simple regression problem with $k = 2$,

$$Y = \beta_1 x_1 + \beta_2 x_2 + e,$$

and the contribution to the inverse of the covariance matrix of $\hat{\underline{\beta}}$ based on a single observation at (x_1, x_2) is given by

$$\frac{1}{\nu^2}\begin{pmatrix} x_1^2 & x_1 x_2 \\ x_1 x_2 & x_2^2 \end{pmatrix}.$$

With the identification

$$x_1 = \frac{1}{\sigma}\psi^{\frac{1}{2}}(z),$$

$$x_2 = \frac{1}{\sigma}z\psi^{\frac{1}{2}}(z),$$

where $z = (x - \mu)/\sigma$, $I(\underline{\theta})$ in (1.1) is of exactly this form and the design problem of minimizing the asymptotic variance of the MLE of $\phi = a_1\mu + a_2\sigma$

can be viewed as the regression problem with

$$S = \left\{ \underline{x} = (x_1, x_2) : -\infty < z < +\infty \right\}.$$

To estimate $\mu = ED50$, let $\underline{a}^T = (1,0)$. The solution shown in Figure 1 is to place all observations at the point $z = 0$, (i.e. $x = \mu$); the corresponding asymptotic variance of $\hat{\mu}$ would be

$$V(\hat{\mu}) = \frac{4\sigma^2}{n}.$$

For estimation of the ED50, exactly the same design is obtained for the probit function; see Chernoff[5].

Of course μ is unknown, but the situation might arise where there is some idea of the value of μ. If this is the case, the above solution would suggest putting all the observations at this suspected value of μ. This design would not be used in practice, however, since even if the main interest is the estimation of the ED50, examination of the fit of the model would also be of importance and this design does not allow for any such test. However it does give an optimal design with which to compare more realistic designs.

3. Alternate Multi-point Designs

In view of the form of the optimal design it seems reasonable to add more dose levels while keeping a large number of observations in the vicinity of the suspected value of μ. Smith, Savin, and Robertson[13] looked at the maximum likelihood estimates of the ED50, for the logit model, and their rate of

convergence to normality, for some 5 and 8 point designs. Their main conclusion was: when inference about the ED50 is of main interest, symmetric (about μ) designs are advisable and extreme response probabilities should be avoided. In view of this work, this section considers some symmetric (about μ) multi-point designs and compares them to the optimal design for estimation of the ED50. Table I shows a number of such 3 point designs and the resulting asymptotic variances of $\hat{\mu}$ for the logit model. $P(x) = \Psi(\frac{x-\mu}{\sigma})$ is assumed known at 3 points, and λ is the fraction of observations at each of dose levels x_1 and x_3, where $P(x_1) = 1 - P(x_3)$ is given at the top of the table. The remaining observations are assumed to be taken at $x_2 = \mu$, where $P(x_2) = \Psi(0) = 0.5$. * From Table I it can be seen that, depending upon the range of $P(x)$ around $P(\mu) = 0.5$ in which one is interested, relatively high efficiency can be acheived. These 3 point designs yield at least an heuristic check on the fit of the model in this range (could at least check assumed symmetry, for example). With an increased range of interest the fraction of observations one must put at μ to acheive the same efficiency increases: 1) for fixed $P(x)$, efficiency decreases as λ increases; and 2) for fixed λ, efficiency decreases as $P(x)$ increases. Table II shows a similar relationship for 5 point symmetric designs. Here λ_1 is the fraction of the n observations put at both x_1 and x_5, with $P(x_1) = 1 - P(x_5)$. Similarily, λ_2 is the fraction of

* Depending on n, the total number of observations, these designs could lead to non-integer allocations.

the n observations put at both x_2 and x_4, with $P(x_2) = 1 - P(x_4)$. $P(x_1)$ and $P(x_2)$ are given at the top of each table. The remaining observations are placed at $x_3 = \mu$ where $P(x_3) = \Psi(0) = 0.5$.

From Table I and Table II it seems reasonable to assume that, if the main interest is estimation of the ED50, and goodness-of-fit of the model is only of interest in a moderate region of $P(x)$ about $P(\mu) = 0.5$ (say $0.2 < P(x) < 0.8$), then given a good initial guess of μ and σ, a pyramid type design, symmetric about μ and within the region of interest, will yield high efficiency. This type of design also allows some assessment of goodness-of-fit of the model.

Of course the guessed values of μ and σ used to design the experiment could be quite poor. Of interest then is the robustness of this type of design to poor initial guesses of μ and σ. To address this question assume that σ is known but our value of μ is incorrect. Suppose the experiment is designed assuming $P(x_0) = 0.5$, but $P(x_0)$ actually equals p^*. This implies

$$\Psi\left(\frac{x_0 - \mu}{\sigma}\right) = p^*,$$

or

$$\mu = x_0 - \sigma\Psi^{-1}(p^*) = x_0 - \sigma \log \frac{p^*}{1 - p^*}.$$

So for any dose level x the actual value of $P(x)$ is

$$\Psi\left(\frac{x - x_0}{\sigma} + \gamma\right),$$

where $\gamma = \log\{p^*/(1-p^*)\}$, but the design assumes:

$$P(x) = \Psi\left(\frac{x - x_0}{\sigma}\right).$$

From this the actual asymptotic variances of the designs in Table I and Table II can be calculated given any specific guessed values of μ. Table III gives the resulting efficiency (to the optimal) of the 3 point design from Table I, with $P(x_1) = 0.2$, for incorrect values of μ. The experiment was designed assuming $P(x_0) = 0.5$. The actual value of $P(x_0)$ is given at the top of each sub-table. Note that as λ approaches 0 the $V(\hat{\mu})$ approaches $+\infty$ for any incorrect guess of μ. Table IV gives the same results for the 5 point design from Table II, with $P(x_1) = 0.2$ and $P(x_2) = 0.4$. These two designs could be thought of as competitors in the situation where goodness-of-fit is only of interest in the range $0.2 < P(x) < 0.8$. The tabulations show that the pyramid designs have higher efficiency if the guessed value of μ is reasonably good, but the equal allocation designs are more robust to poor guesses of μ.

The question of robustness can also be addressed without assuming σ is known, but the evaluations become less decipherable. It should be noted that incorrect guesses of μ affect the symmetry of the design, while incorrect guesses of σ affect the distance from μ at which the design points are taken. So, though all-encompassing statements are difficult, some idea of the relative importance of the accuracy of guesses of μ and σ can be obtained by comparing the effect of poor guesses of μ, outlined in Table III, and the

effect of decreasing $P(x_1)$ in Table I.

4. An Example

The Department of Fisheries and Oceans, Vancouver BC recently sponsored a survey of sport fishing in British Columbia. * As part of the analysis of the data collected, the logit quantal response model was used to estimate the economic value of sport fishing in BC tidal waters. Along with some background information, four questions were asked of fishermen returning to docks in four major fishing areas on Vancouver Island:

1) How many days do you plan to go fishing between now and the end of the next month? $\underline{D}$ days

2) Suppose you were offered $\underline{D}$ (days) x $\underline{y} = \underline{D}y$ dollars to give up fishing in tidal waters until the end of next month. Would you accept the offer? No___ Yes___

3)How much did you spend on your fishing trip today? (Include costs such as, bait, gasoline, boat rentals. Do not include equipment costing more than \$100) \$_____

4)Now imagine that the cost of fishing in BC tidal waters increased. If the cost of your fishing trip had been $\underline{z}$ dollars higher today would you still have gone fishing? No___ Yes___

* Economic Valuation of the BC Tidal Sport Fishery. DPA Group Inc. Vancouver BC, February 1985.

In question 2 the amount offered, y, can be viewed as the dose level with the answer a binary response. Similarily in question 4, the increased fishing cost, z, can be viewed as a dose level with the answer a binary response. Of main interest in this study was the ED50 for each of the four geographical areas, since the ED50 represents an estimate of the net "average value" per angler day of the sport fishing experience. The Department of Fisheries and Oceans set 30 dose levels for each question ranging from \$2 to \$80 per day for question 2 and \$1 to \$50 for question 4, and specified a target of an equal number of observations at each dose level. As the survey continued it became apparent the dose levels chosen were a bit low in terms of symmetry about the ED50; hence, the range of dose levels was extended for the second half of the study. For question 2 the dose range was extended to include \$2 to \$200 and for question 4 to include \$1 to \$100 with equal allocation at each dose level from this point on. Thus overall there were approximately twice as many observations on the dose levels in the initial range as the added ones. Question 4 for the Sechelt area will be used for comparisons of this design to the designs of Section 2. The maximum likelihood estimates, $\hat{\mu}$ and $\hat{\sigma}$, for μ and σ , and the variance of $\hat{\mu}$ obtained in this study, based on the $n = 382$ observations are as follows:

$$\hat{\mu} = 41.34, \ \hat{\sigma} = -13.05$$

$$\mathrm{V}(\hat{\mu}) = 7.40.$$

Due to the nature of the question, the higher the dose the fewer the number of positive responses, and thus the negative value of σ.

If we assume the estimated values are in fact the exact values of μ and σ, then comparisons can be made. If the 5 point design of Table IV ($P(x_1) = 0.2$, $P(x_2) = 0.4$) had been used with $\lambda_1 = 0.05$ and $\lambda_2 = 0.2$, the estimate of μ obtained would have had asymptotic variance less than above, as long as the initial guess μ_0 of μ was in the range $31 \leq \mu_0 \leq 51.5$, assuming σ had been guessed correctly. If the correct value of μ and σ had been guessed the attainable $V(\hat{\mu})$ with this 5 point design would have been (see Table II):

$$V(\hat{\mu}) = 4.22 \cdot (13.05)^2/382 = 1.88,$$

and $n = 98$ would have been sufficient to attain the asymptotic variance of 7.40 achieved in the study. Figure 2, a graph of $V(\hat{\mu})$ vs μ_0, shows explicitly the relationship between μ_0 and the achievable variance using this 5 point design and the variance obtained in the study (7.40). Table V gives $\hat{\mu}$, $\hat{\sigma}$, and $V(\hat{\mu})$ obtained in the study, as well as the attainable asymptotic $V(\hat{\mu})$ using the above 5 point design with correct guesses of μ and σ for question 2 and 4 at each of the geographical regions considered.

In this survey it is apparent that no accurate knowledge of μ and σ was available prior to designing the experiment. However the use of such a large number of dose levels does not seem to be warranted. In addition to the loss of efficiency which has been demonstrated, the use of a large number

of dose levels must have greatly complicated the study due to the need for randomization and balancing over both time and regions. It should also be noted that, in this study some of the dose levels used obviously implied a priori extreme probabilities of response. The extreme dose levels close to zero provided essentially no information about the ED50. It would seem that more realistic lower bounds could have been possible. As noted in the previous section, numerical work suggests that this could affect the asymptotics used in inference about the ED50.

The initial design of equal allocation to 30 dose levels between \$2 and \$80 (for question 2) seems to indicate a willingness to assume \$2 to \$80 encompasses a reasonable range of $P(x)$ about the ED50, but with no committment to any specific point within this range as an initial guess of μ. A 5 point design with equal allocation at dose levels equally spaced between \$2 and \$80 would have competitive variance for any value of μ, and represents relatively the same amount of prior knowledge while greatly simplifying the survey. To illustrate this Table VI compares the attainable variance of $\hat{\mu}$ using equal allocation to the 30 dose levels used initially in the study to the attainable variance of $\hat{\mu}$ using equal allocation to the 5 points (\$8, \$24, \$40, \$56, \$72) for various values of μ, and assuming $\sigma = -13.05$, the value obtained for Sechelt question 4. It should also be noted that, when it became apparent in the study that the choice of dose levels was too low, and that more were to be

added, a sequential approach could have been used to choose the next set of dose levels. A reasonable approach would be to estimate μ and σ using the information already obtained, and design the second half of the experiment assuming these are the true values of μ and σ (see Wetherill[14], for example).

II. Alternate Criteria For Design Of Quantal Response Experiments

1. Introduction to the Problem

When designing a quantal response experiment where estimation of μ, the ED50, is of main interest, minimizing the asymptotic variance of $\hat{\mu}$, the MLE of μ, with respect to dose allocation and number of doses, given a good initial guess $\underline{\theta}_0 = (\mu_0, \sigma_0)$ of $\underline{\theta} = (\mu, \sigma)$, leads to an inadvisable design as described in chapter I. One would like some alternate criterion which will assure some ability to test the goodness-of-fit of the presumed underlying model. Finney[7] proposed the criterion of minimizing the square of the half-length of a fiducial interval for μ, given a good initial guess μ_0 of μ. Kalish and Rosenberger[8] consider 2 point designs symmetric about $x = \mu$, and determine D-optimal, G-optimal, A-optimal, and E-optimal designs. A design which minimizes the determinant of $I^{-1}(\underline{\theta})$, where $I(\underline{\theta})$ is the total information matrix for $\underline{\theta}$, is called D-optimal, and a design which minimizes the maximum variance of a predicted response over a specific region of the explanatory variables is called G-optimal. A-optimality refers to minimizing the trace of $I^{-1}(\underline{\theta})$, and E-optimality to minimizing the maximum latent root. All of these criterion need a good initial guess of $\underline{\theta}$, since $I(\underline{\theta})$ depends on $\underline{\theta}$.

2. Alternate Criterion 1

When modelling a quantal response experiment as described in chapter I, the probability of response $p(x)$ at dose level x is assumed to be $F(x \mid \underline{\theta})$ for a specified cumulative distribution function F. The maximum likelihood estimate $\hat{\underline{\theta}}$ of $\underline{\theta}$ can then be obtained, as well as an estimate of Σ, the covariance matrix for $\hat{\underline{\theta}}$. Using $\hat{\underline{\theta}}$, the MLE for $p(x) = F(x \mid \underline{\theta})$ for a given x is $\hat{p}(x) = F(x \mid \hat{\underline{\theta}})$. A possible criterion for designing such an experiment under the model assumption is to choose the design D to minimize the expected overall distance between $F(x \mid \underline{\theta})$ and $F(x \mid \hat{\underline{\theta}})$:

$$\min_D \mathrm{E}_{\hat{\underline{\theta}}}\left\{\int_{-\infty}^{+\infty}\left[F(x \mid \hat{\underline{\theta}}) - F(x \mid \underline{\theta})\right]^2 dx\right\}.$$

The integral over dose gives an overall measure of the distance between the $\hat{p}(x)$ curve and the true $p(x)$ curve; Kuo[9] uses (a weighted version of) this distance measure as the loss function in a Bayesian nonparametric approach to the same problem. The criterion then suggests minimizing the expected value of this *distance* over all possible designs D.

First

$$C(D) = \mathrm{E}_{\hat{\underline{\theta}}}\left\{\int_{-\infty}^{+\infty}\left[F(x \mid \hat{\underline{\theta}}) - F(x \mid \underline{\theta})\right]^2 dx\right\}$$

must be evaluated. The asymptotic value can be obtained as a function of Σ, the covariance matrix of $\hat{\underline{\theta}}$. Assuming $F(x \mid \hat{\underline{\theta}})$ can be expanded in a Taylor series about $F(x \mid \underline{\theta})$ yields

$$F(x \mid \hat{\underline{\theta}}) \doteq F(x \mid \underline{\theta}) + (\hat{\underline{\theta}} - \underline{\theta})^T\left(\frac{\partial}{\partial \underline{\theta}} F(x \mid \underline{\theta})\right),$$

which implies

$$C(D) \doteq \mathrm{E}_{\underline{\hat{\theta}}}\left\{\int_{-\infty}^{+\infty}\left[(\underline{\hat{\theta}}-\underline{\theta})^T\left(\frac{\partial}{\partial\underline{\theta}}F(x\mid\underline{\theta})\right)\right]^2 dx\right\}.$$

In matrix notation this becomes

$$\begin{aligned} C(D) &\doteq \mathrm{E}_{\underline{\hat{\theta}}}\left\{\int_{-\infty}^{+\infty}\left(\frac{\partial}{\partial\underline{\theta}}F(x\mid\underline{\theta})\right)^T(\underline{\hat{\theta}}-\underline{\theta})(\underline{\hat{\theta}}-\underline{\theta})^T\left(\frac{\partial}{\partial\underline{\theta}}F(x\mid\underline{\theta})\right)dx\right\} \\ &= \int_{-\infty}^{+\infty}\left(\frac{\partial}{\partial\underline{\theta}}F(x\mid\underline{\theta})\right)^T \mathrm{E}_{\underline{\hat{\theta}}}\left[(\underline{\hat{\theta}}-\underline{\theta})(\underline{\hat{\theta}}-\underline{\theta})^T\right]\left(\frac{\partial}{\partial\underline{\theta}}F(x\mid\underline{\theta})\right)dx. \end{aligned}$$

This implies

$$C(D) \doteq \int_{-\infty}^{+\infty}\left(\frac{\partial}{\partial\underline{\theta}}F(x\mid\underline{\theta})\right)^T\Sigma\left(\frac{\partial}{\partial\underline{\theta}}F(x\mid\underline{\theta})\right)dx \tag{2.1}$$

since $\mathrm{E}(\underline{\hat{\theta}}) \doteq \underline{\theta}$.

For a location-scale model, $\underline{\theta} = (\mu,\sigma)^T$, where $F(x \mid \underline{\theta}) = H\left(\frac{x-\mu}{\sigma}\right)$ for some cumulative distribution function H with density h. In this situation

$$\frac{\partial}{\partial\mu}F = -\frac{1}{\sigma}h\left(\frac{x-\mu}{\sigma}\right)$$

and

$$\frac{\partial}{\partial\sigma}F = -\frac{1}{\sigma}\left(\frac{x-\mu}{\sigma}\right)h\left(\frac{x-\mu}{\sigma}\right).$$

Substituting these into equation (2.1) yields

$$\begin{aligned} C(D) \doteq \frac{1}{\sigma}\Bigg\{ &\mathrm{V}(\hat{\mu})\int_{-\infty}^{+\infty} h^2(t)dt + \mathrm{V}(\hat{\sigma})\int_{-\infty}^{+\infty} t^2\, h^2(t)dt \\ &+ 2\mathrm{Cov}(\hat{\mu},\hat{\sigma})\int_{-\infty}^{+\infty} t\, h^2(t)dt\Bigg\}. \end{aligned} \tag{2.2}$$

Further, if $h(\cdot)$ is symmetric about 0 then

$$\int_{-\infty}^{+\infty} t\, h^2(t)dt = 0,$$

in which case (2.2) reduces to

$$C(D) \doteq \frac{1}{\sigma}\left\{ \mathrm{V}(\hat{\mu}) \int_{-\infty}^{+\infty} h^2(t)dt + \mathrm{V}(\hat{\sigma}) \int_{-\infty}^{+\infty} t^2 h^2(t)dt \right\}. \qquad (2.3)$$

Finally, for the logit model ($H = \Psi, h = \psi$)

$$\int_{-\infty}^{+\infty} h^2(t)dt = \frac{1}{6}$$

and

$$\int_{-\infty}^{+\infty} t^2\, h^2(t)dt = \frac{\pi^2 - 6}{18}.$$

Thus for the logit model (2.3) reduces to

$$C(D) \doteq \frac{1}{\sigma}\left\{ \frac{1}{6}\mathrm{V}(\hat{\mu}) + \left(\frac{\pi^2 - 6}{18}\right)\mathrm{V}(\hat{\sigma}) \right\}, \qquad (2.4)$$

where $\mathrm{V}(\hat{\mu}) = I_{11}^{-1}(\underline{\theta})$, and $\mathrm{V}(\hat{\sigma}) = I_{22}^{-1}(\underline{\theta})$ obtained from $I(\underline{\theta})$ in (1.1). The proposal is to minimize $C(D)$ as approximated in (2.4) with respect to $\lambda_i = n_i/n$, x_i, and k, where $i = 1, ..., k$. This was done numerically (using a successively refined grid search) for fixed values of k assuming good initial guesses μ_0, of μ. For $k = 2$ the design obtained consisted of: x_1 and x_2 chosen such that $\mathrm{P}(x_1) = 1 - \mathrm{P}(x_2) \doteq 0.2$ with $\lambda_1 = \lambda_2 = \frac{1}{2}$. Setting $k = 3$ and $k = 5$ did not change the design obtained.

As described in Chapter I, Figure 1 provides the basis for determining the optimal design for the minimization of the asymptotic variance of any linear combination of $\hat{\mu}$ and $\hat{\sigma}$, $\mathrm{V}(a\hat{\mu} + b\hat{\sigma})$, where a and b are specified constants; these designs are 1 or 2 point designs depending on the particular linear

combination of interest. Here we wish to minimize $aV(\hat{\mu}) + bV(\hat{\sigma})$ and the covariance term which plays an important role in the above problem does not enter. Nevertheless, the situation is similar and it is not surprising that a 2 point design is advised.

Another way of obtaining a design using the $C(D)$ criterion would be to choose the design to minimize the asymptotic variance of $\hat{\mu}$ subject to $n\sigma C(D) \leq Q$, where Q is some specified constant. There does not appear to be any natural way of specifying this constant however, and this approach is not pursued further.

3. Alternate Criterion 2

Criterion 1 proposed in the previous section is based on minimizing a measure of the expected distance between the actual function, $p(x)$, and the MLE, $\hat{p}(x)$, of $p(x)$ assuming the model is correct. The design problem can be viewed in an alternate way. The experiment may be designed to ensure a powerful test against some specific deviation from the assumed model. Chapman and Nam[4] discuss this criterion for the case

$$p(x) = \alpha + \beta x.$$

The discussion is based on the Pearson's chi-square test associated with the hypothesis of interest. A formulation of some asymptotic results used by Chapman and Nam[4], and general enough for this situation, follows.

An experiment consists of k sequences of n_i trials ($\sum_{i=1}^{k} n_i = n$), where

each trial may result in an event E or its complement $\bar{E}$. Let $p_i = p_i(\underline{\theta})$ be the probability of event E on the i^{th} sequence, where $\underline{\theta} = (\theta_1, ..., \theta_m)^T$ belongs to a set M in m-dimensional Euclidean space (assume $m < k$). Let $p_i^0 = p_i(\underline{\theta}^0)$, where $\underline{\theta}^0$ is an interior point of M. Let u_i be the number of occurences of E on the i^{th} sequence. With suitable conditions on the set M and the functions $p_i(\underline{\theta})$, under the sequence $(n \to +\infty)$ of alternatives

$$H_A : p_i = p_i^0 + c_i/\sqrt{n},$$

the statistic

$$X^2 = \sum_{i=1}^{k} \frac{\left[u_i - n_i p_i(\hat{\underline{\theta}})\right]^2}{n_i p_i(\hat{\underline{\theta}})\left(1 - p_i(\hat{\underline{\theta}})\right)},$$

where $\hat{\underline{\theta}}$ is an asymptotically efficient estimator, is asymptotically, for $\lambda_i = n_i/n$ fixed and as $n \to +\infty$, distributed as $\chi^2_{[k-m]}(\Delta)$; the noncentrality parameter Δ is given by (see Mitra[10])

$$\Delta = \delta^T \left[I - B(B^T B)^{-1} B^T\right]\delta, \tag{2.5}$$

where

$$\delta^T = \left[\frac{c_i\sqrt{\lambda_i}}{\sqrt{p_i^0}}, \frac{-c_i\sqrt{\lambda_i}}{\sqrt{1-p_i^0}} : i = 1...k\right]_{1\times 2k}, \tag{2.6}$$

and

$$B^T = \left[\frac{\sqrt{\lambda_i}}{\sqrt{p_i^0}}\left(\frac{\partial}{\partial\theta_h}p_i\right)_{\theta_h=\theta_h^0}, \frac{\sqrt{\lambda_i}}{\sqrt{1-p_i^0}}\left(\frac{\partial}{\partial\theta_h}(1-p_i)\right)_{\theta_h=\theta_h^0} : i = 1...k\right]_{m\times 2k}. \tag{2.7}$$

In the dose response problem E represents a response, and $\bar{E}$ represents no response. The experiment can be constructed to ensure a powerful test

of the hypothesis

$$H_0 : p_i = F(x_i \mid \underline{\theta}^0)$$

against

$$H_A : p_i = F(x_i \mid \underline{\theta}^0) + c_i/\sqrt{n}$$

The unknown vector of parameters $\underline{\theta}$ can be estimated using maximum likelihood estimation. The choice of the c_i's depends upon the deviation from the model which is of most interest.

For the logit model $F(x \mid \underline{\theta}) = \Psi\left(\frac{x-\mu}{\sigma}\right)$ and

$$X^2 = \sum_{i=1}^{k} \frac{\left[u_i - n_i\Psi(\hat{z}_i)\right]^2}{n_i\psi(\hat{z}_i)}$$

where $\hat{z}_i = (x_i - \hat{\mu})/\hat{\sigma}$, $\Psi(t) = e^t/(1+e^t)$ and $\psi(t) = e^t/(1+e^t)^2$ as in chapter I. The general theory implies that, in the limit, X^2 will have a chi-square distribution with $k-2$ degrees of freedom under H_0, and a non-central chi-square distribution with $k-2$ degrees of freedom and non-centrality parameter Δ under H_A. From (2.6) and (2.7)

$$\delta^T B = \left[-\frac{1}{\sigma_0}\sum_{i=1}^{k} c_i\lambda_i, -\frac{1}{\sigma_0}\sum_{i=1}^{k} c_i\lambda_i z_i^0\right] \tag{2.8}$$

and

$$B^T B = \frac{1}{\sigma^2}\begin{bmatrix} A & C \\ C & D \end{bmatrix} \tag{2.9}$$

where $A = \sum_{i=1}^{k} \lambda_i\psi(z_i^0)$, $C = \sum_{i=1}^{k} \lambda_i z_i^0\psi(z_i^0)$, $D = \sum_{i=1}^{k} \lambda_i(z_i^0)^2\psi(z_i^0)$, and $z_i^0 = (x_i - \mu_0)/\sigma_0$. Thus in this case $B^T B = \frac{1}{n}I(\underline{\theta}^0)$, where $I(\cdot)$ is given

in (1.1), and Δ can then be written in terms of (2.8) and (2.9) as follows:

$$\Delta = \delta^T \delta - (\delta^T B)(B^T B)^{-1}(\delta^T B)^T.$$

Choosing the design to minimize the asymptotic variance $\hat{\mu}$ under the restriction $\Delta > \Delta_0$, where Δ_0 is some specified constant, might appear to be a desirable way to proceed. But the above development is for a particular sequence of alternatives. The noncentrality parameter, Δ, is a function of the particular value of the parameter $\underline{\theta}^0$ through $z_i^0 = (x_i - \mu_0)/\sigma_0$, and we only have the noncentrality parameter for all $\underline{p}$ "close enough" to $F(x_i \mid \underline{\theta}^0)$. This is the usual type of situation: if we want to discuss power, we have to be willing (able) to specify the magnitude of the departures of interest. But the situation in our design problem is more complex; the objective is to ensure that designs prescribed will be reasonably sensitive for detecting departures from the assumed underlying model logit $p(x) = (x - \mu)/\sigma$, where the values of the parameters are not specified. Also note that the c_i's which appear in the noncentrality parameter are deviations at particular x_i's, but these x_i's are part of what is determined in the design problem. Given $\{x_i\}_{i=1}^k$, $\underline{\theta}_0$, and $\{c_i\}_{i=1}^k$, one could determine how the observations should be allocated to maximize Δ, and also what allocations yield $\Delta \geq \Delta_0$, but this does not really address the design problem of interest.

III. Quantal Response Experiments Over Time

1. Description of Problem

In chapter I a typical quantal response problem and the standard type of analysis for such problems were described. In this chapter a generalization of this problem is examined, and a method of analysis is proposed. In the generalized problem subjects are assigned to one of k fixed dose levels. At each of m fixed time points the subjects are classified into one of two categories, has responded or has not responded, and the number of subjects that have responded is noted. This is done for each dose level separately. A number of replications of this basic experiment could be performed.

Carter and Hubert[3] proposed the following growth-curve model approach to such problems.

notation

d_i	the ith dose level, $i = 1, ..., k$;
x_i	$= \log_{10}(d_i)$;
t_j	the jth time point, $j = 1, ..., m$;
n_{il}	the number of subjects at dose level d_i in replication l, $l = 1, ..., L$;
r_{ijl}	the number of subjects at dose level d_i in replication l that respond prior to time t_j.

First a transformation of the raw data was carried out:

$$z_{ijl} = \arcsin(r_{ijl}/n_{il})^{\frac{1}{2}}.$$

Then the response variables Z_{ijl} were modelled using a polynomial of order

$q-1$ in time and linear in $\log_{10}$(dose) as follows:

$$Z_{ijl} = \sum_{s=1}^{q} (\beta_{s1} + \beta_{s2} x_i + \rho_{sl}) t_j^{-(s-1)} + \epsilon_{ijl}$$

where $\sum_{l=1}^{L} \rho_{sl} = 0$ for each s. The error vectors

$$\underline{\epsilon}_{il} = (\epsilon_{i1l}, \epsilon_{i2l}, ..., \epsilon_{iml})^T,$$

are assumed to be independent and normally distributed with mean 0 and completely arbitrary (unknown) covariance matrix Σ. With the obvious vector notation, this model can be expressed as: $\underline{Z}_{il}$ are independent random vectors with

$$\underline{Z}_{il} \sim N_m(\underline{\mu}_{il}, \Sigma)$$

where

$$\mu_{ijl} = \sum_{s=1}^{q} (\beta_{s1} + \beta_{s2} x_i + \rho_{sl}) t_j^{-(s-1)}.$$

Under the stated assumptions, this model can then be analyzed using the growth curve methodology of Potthoff and Roy[12]. This has the advantage that the maximum likelihood estimators of $\underline{\beta} = (\beta_{s1}, \beta_{s2} : s = 1, ..., q)$, and $\underline{\rho} = (\rho_{sl} : s = 1, ..., q;\ l = 1, ..., L)$ have closed form solutions, and under the model assumptions all the necessary distribution theory is available. This allows straightforward determination of confidence intervals for the ED50 at any fixed time point, as well as simultaneous confidence bands for the ED50 as a function of time, without appealing to asymptotic results. Carter and Hubert[3] present the results of an application of this model to a data set

consisting of $L = 2$ replications of a basic experiment involving $n_{il} = 10$ fish on each of $k = 7$ dose groups, all observed at the same $m = 3$ time points. A particular feature of the experiment is that the 10 fish on a particular dose group are all contained in a single tank to which a specified concentration of a toxic copper substance was administered; thus the experimental units are the tanks, while the sampling units are the fish.

There appear to be some potential difficulties with the use of this model in the present context. The first potential difficulty involves the normality assumption. For fixed i and l, the possible values of $R_{ijl}(j = 1, ..., m)$ are $0, 1, ..., n_{il}$, but these response variables are further restricted by the relationship $R_{i1l} \leq R_{i2l} \leq \cdots \leq R_{iml}$. Since $\arcsin(\cdot)^{\frac{1}{2}}$ is a monotonic function, the same restriction applies to the transformed response variables $Z_{ijl}(j = 1, ..., m)$. This immediately raises the issue of the adequacy of the multivariate normal approximation for the distribution of $\underline{Z}_{il}$. While any parametric analysis will involve some distributional approximation, the use of the multivariate normal seems particularly questionable in the application presented by Carter and Hubert[3] where $n_{il} = 10$ (and $m = 3$). Here a trivariate distribution restricted as described above and with each univariate marginal distribution supported on the set of points $\left\{ \arcsin(r/10)^{\frac{1}{2}} : r = 0, 1, ..., 10 \right\}$ is being approximated by a trivariate normal distribution; the adequacy of this approximation would appear doubtful to say the least.

The second potential difficulty involves the assumption concerning the covariance structure; the $\underline{Z}_{il}$ are assumed to be independent and normally distributed with unknown covariance matrix $\Sigma_{il} = \Sigma$, where Σ is of completely general structure, and does not vary with dose or replication. It is unclear how such an assumption can be justified; the nature of the response variables provides some information concerning the approximate structure of Σ_{il}, and clearly suggests that Σ_{il} should be allowed to vary from dose to dose within each replication. To see this, consider the data for an individual subject. After the dose is administered to the subject at time $t_0 = 0$, the subject remains in the no-response category until his (random) time of response T, when he becomes a member of the response category where he remains from time T onward. Each subject is observed at the same m time points $t_1, ..., t_m$ thereby identifying the interval in which he responded. Suppose the response times for the subjects on dose group i and replication l are assumed to be independent and identically distributed according to the response time distribution F_{il}:

$$F_{il}(t) = \mathrm{P}(T_{il} \leq t),$$

where T_{il} represents the response time for any one of these subjects. Let

$$p_{ijl} = \mathrm{P}(t_{j-1} < T_{il} \leq t_j) = F_{il}(t_j) - F_{il}(t_{j-1})$$

for $j = 1, ..., m+1$ with $t_{m+1} = +\infty$, and note that $\sum_{\nu=1}^{j} p_{i\nu l} = F_{il}(t_j)$. If U_{ijl} denotes the number of subjects at dose level i and replication l who re-

spond between t_{j-1} and t_j, then the random vector $\underline{U}_{il} = (U_{il1}, ..., U_{i,m+1,l})^T$ is multinomially distributed with index n_{il} and cell probabilities p_{ijl} $(j = 1, ..., m+1)$ for each $i = 1, ..., k$ and $l = 1, ..., L$. The data under consideration, r_{ijl}, is an observation of the cumulative count $R_{ijl} = \sum_{\nu=1}^{j} U_{i\nu l}$. It follows that

$$\mathrm{Cov}(R_{ijl}, R_{ij'l}) = n_{il}\Big(\sum_{\nu=1}^{j} p_{i\nu l}\Big)\Big(1 - \sum_{\nu=1}^{j'} p_{i\nu l}\Big) = n_{il}F_{il}(t_j)\Big[1 - F_{il}(t_{j'})\Big],$$

for $j' \geq j$. As the first step of their approach, Carter and Hubert[3] perform what they refer to as a variance-stabilizing transformation:

$$Z_{ijl} = \arcsin(R_{ijl}/n_{il})^{\frac{1}{2}}.$$

Under the above assumptions, the delta method yields the approximate covariance structure of $\underline{Z}_{il}$ as

$$\begin{aligned}\mathrm{Cov}(Z_{ijl}, Z_{ij'l}) &\doteq \frac{1}{4n_{il}}\sqrt{\frac{\Big(\sum_{\nu=1}^{j} p_{i\nu l}\Big)\Big(1 - \sum_{\nu=1}^{j'} p_{i\nu l}\Big)}{\Big(\sum_{\nu=1}^{j'} p_{i\nu l}\Big)\Big(1 - \sum_{\nu=1}^{j} p_{i\nu l}\Big)}} \\ &= \frac{1}{4n_{il}}\sqrt{\frac{F_{il}(t_j)\Big[1 - F_{il}(t_{j'})\Big]}{F_{il}(t_{j'})\Big[1 - F_{il}(t_j)\Big]}},\end{aligned}$$

for $j' \geq j$. When $j' = j$ this becomes

$$\mathrm{V}(Z_{ijl}) \doteq \frac{1}{4n_{il}}.$$

It can be seen that at a given dose and replication

$$\Sigma_{il} = \Big(\mathrm{Cov}(Z_{ijl}, Z_{ij'l}) : j = 1, ..., m; j' = 1, ..., m\Big)_{m \times m}$$

will have more weight on the diagonal with weight decreasing away from the diagonal. Also Σ_{il} depends upon both $\underline{p}_{il} = (p_{i1l}, ..., p_{i,m+1,l})$ and n_{il}, so even in the case where the n_{il} are all equal, Σ_{il} will vary across doses and replications if, as is anticipated, the $\underline{p}_{il}$ do.

Carter and Hubert[3] emphasize that their approach is intended to differentiate between sampling units and experimental units. In the context of their application, they want to allow for the possibility that there may be a cause of mortality other than the toxic copper substance that affects all of the fish in a tank simultaneously. In this case, there are two sources of variation in the mortality counts corresponding to the different time intervals; multinomial variation affecting each fish and tank variation affecting all of the fish in a tank. Their assumption of a common unknown covariance matrix Σ for the vectors $\underline{Z}_{il}$ of transformed cumulative counts appears to be an attempt to account for the unknown tank variation.

While the above discussion does not differentiate between sampling and experimental units, this additional source of variation could easily and explicitly be incorporated by a slight extension of the model. Possibly the simplest way to do so would be to assume that the vector of cell probabilities $\underline{p}_{il}$ corresponding to the i^{th} dose group in replication l is itself randomly distributed across tanks. If this distribution is taken to be the Dirichlet

distribution defined by the density

$$h(\underline{p}_{il}) = \Gamma(\omega) \prod_{j=1}^{m+1} \frac{p_{ijl}^{\omega\pi_{ijl}-1}}{\Gamma(\omega\pi_{ijl})},$$

where $\omega > 0$, $\pi_{ijl} > 0$ and $\sum_{j=1}^{m+1} \pi_{ijl} = 1$, then the unconditional distribution of $\underline{U}_{il}$, the vector of mortality counts, is the Dirichlet-multinomial; see Moismann[11]. Unconditionally, we have

$$\mathrm{E}(\underline{U}_{il}) = n_{il}\underline{\pi}_{il},$$

and

$$\mathrm{V}(\underline{U}_{il}) = n_{il}C_{il}\left(\mathrm{diag}(\underline{\pi}_{il}) - \underline{\pi}_{il}\underline{\pi}_{il}^{T}\right),$$

where $C_{il} = (n_{il} + \omega)/(1 + \omega)$. Thus the covariance matrix is a constant, C_{il}, times the corresponding multinomial covariance matrix based on $\underline{\pi}_{il}$. It follows that except for multiplication by the variance inflation factor C_{il}, the covariance structure of $\underline{Z}_{il}$ is of the same form as described above.

There are alternate methods of incorporating this additional source of variation, but explicit assumptions leading to the result that the $\underline{Z}_{il}$ have common (completely arbitrary) unknown covariance matrix Σ seem difficult to identify. It appears this assumption has been made as a matter of convenience so that the observed data exhibits the probabilistic structure required for the application of the Potthoff-Roy growth-curve methodology.

An alternate, more direct approach, which would not require any assumptions which are clearly incorrect at the outset, would involve an analysis based on the multinomial structure of the vector of mortality counts

$\underline{U}_{il}$, together with a parametric specification of the underlying response time distribution F_{il}. Such analyses do not incorporate any tank variation which may be present; whether such a generalization is required can be subsequently examined via the use of the Dirichlet-multinomial model in place of the multinomial model. Various aspects of statistical inference for the Dirichlet-multinomial model, which are relevant to such an undertaking, are considered in Brier[2] and Wilson[15]. An alternate analysis along these lines, of the data set presented by Carter and Hubert[3], will be pursued in the remainder of this chapter.

2. Proposed Model

To motivate a more direct method for analyzing this type of problem, consider the data corresponding to a single dose level within a particular replication: n subjects are treated at dose level d and the cumulative number of responses r_j are observed at times t_j, $j = 1, ..., m$. Assume the subjects respond independently with

$$\mathrm{P}(T \leq t) = F(t)$$

where T is the time to response. If U_j equals the number of responses between t_{j-1} and t_j $(u_j = r_j - r_{j-1})$ then

$$\mathrm{P}(U_1 = u_1, ..., U_{m+1} = u_{m+1}) = n! \prod_{j=1}^{m+1} \frac{[F(t_j) - F(t_{j-1})]^{u_j}}{u_j!}$$

where $t_0 = 0$, $t_{m+1} = +\infty$, and $\sum_{j=1}^{m+1} u_j = n$.

Randomization of subjects to dose levels reduces the general problem to k independent experiments of this type within each of L independent replications. The dose and replication effects could then be incorporated by allowing $F(t)$ to be different for different dose levels and replications. This is usually done by keeping the form of $F(t)$ fixed, and allowing the parameters in $F(t)$ to depend upon dose and replication. In general let $F(t)$ be of the form $F(t, \underline{\theta}_l)$, where $\underline{\theta}_l$ is a vector of parameters for each l. Then the dependence on dose could be incorporated as follows:

$$\underline{\theta}_l = \underline{\theta}_l(d).$$

To motivate the choice of $F(t, \underline{\theta}_l)$ and $\underline{\theta}_l(d)$ some preliminary data analysis was performed on the experimental data presented by Carter and Hubert[3]. Initially an exponential $F(t) = 1 - \exp(-\lambda t)$ was fit at each dose level within each replication, and a reasonable fit was attained ($X^2 = 25.02$, $G^2 = 22.81$, $df = 28$). The relationship between the $\hat{\lambda}$'s and the dose levels was then examined within each replication. The plots of $\hat{\lambda}$ vs dose, $\hat{\lambda}$ vs log(dose), $\log(\hat{\lambda})$ vs dose, and $\log(\hat{\lambda})$ vs log(dose) all appeared reasonably linear, with those involving log(dose) appearing most nearly so.

This preliminary analysis suggested the exponential distribution might provide an adequate model for the time to response. A generalization is provided by the Weibull distribution, which has been widely used in time-

to-response problems. Thus the simple model proposed is

$$F(t) = 1 - \exp(-\lambda t^{\gamma})$$

where $\lambda > 0$ and $\gamma > 0$, with the dependence on dose and replications incorporated as

$$F(t \mid d, l) = 1 - \exp(-\lambda_l(d) t^{\gamma_l(d)}),$$

where $\lambda_l(d) = exp[\alpha_l + \beta_l \log(d)]$ and $\gamma_l(d) = \gamma_l$, with $-\infty < \alpha_l < +\infty$, $\beta_l > 0$, and $\gamma_l > 0$. Viewed as a function of $x = \log_e(d)$ for fixed t and l, this $F(t \mid x, l)$ has the form of a cumulative distribution function as is usually desired for a dose response problem. This model has a $3L$ component vector of parameters $\underline{\theta} = (\underline{\theta}_l : l = 1, ..., L)^T$ where $\underline{\theta}_l = (\alpha_l, \beta_l, \gamma_l)^T$, and except for an additive constant which does not depend upon $\underline{\theta}$, the log-likelihood function is

$$L(\underline{\theta}) = \sum_{i,j,l} u_{ijl} \log\left[F(t_j \mid x_i, \underline{\theta}_l) - F(t_{j-1} \mid x_i, \underline{\theta}_l)\right]$$

where u_{ijl} is the observed number of responses between t_{j-1} and t_j at dose level i on replication l, and $i = 1, ..., k$, $j = 1, ..., m+1$, $l = 1, ..., L$,

3. Estimation and Confidence Intervals for the ED50

The maximum likelihood estimator $\hat{\underline{\theta}}$ for $\underline{\theta}$ can be obtained using an iterative method such as the Newton-Raphson method. Provided $n_l = \sum_{i=1}^{k} n_{il}$ is reasonably large for each l, $I^{-1}(\underline{\theta})$, the inverse of the (total) information matrix for $\underline{\theta}$, will be a good estimate of the covariance matrix for $\hat{\underline{\theta}}$. Although

$I(\underline{\theta})$ depends upon $\underline{\theta}$ which is unknown, $I(\hat{\underline{\theta}})$ can be used as an estimate of $I(\underline{\theta})$. The observed information matrix

$$I_0(\hat{\underline{\theta}}) = \left(-\frac{\partial^2}{\partial\theta_i \partial\theta_j} L(\underline{\theta})\right)_{\underline{\theta}=\hat{\underline{\theta}}}$$

can also be used to estimate $I(\underline{\theta})$, and is more convenient to use since it is evaluated in the course of the Newton-Raphson iteration for $\hat{\underline{\theta}}$. Then the asymptotic normality of maximum likelihood estimators can be used to carry out inference for any of the parameters of interest.

Of particular interest in the present context is inference for the ED50 at a given time point within a particular replication if replication effects are present. Corresponding to the usual situation for quantal response problems, for fixed t within replication l, the log ED50 is the point $x_{0l}(t)$ which satisfies

$$F(t \mid x_{0l}(t), l) = \frac{1}{2}.$$

Under the presumed model, this becomes

$$x_{0l}(t) = \frac{1}{\beta_l}[\log\log(2) - \gamma_l \log(t) - \alpha_l],$$

and the maximum likelihood estimator of $x_{0l}(t)$ is

$$\hat{x}_{0l}(t) = \frac{1}{\hat{\beta}_l}[\log\log(2) - \hat{\gamma}_l \log(t) - \hat{\alpha}_l].$$

Confidence intervals for x_{0l} (the dependence on t is suppressed from here on) can be obtained using one of the following two methods. Method 1 assumes that

$$\hat{x}_{0l} \sim N\big(x_{0l}, \mathrm{V}(\hat{x}_{0l})\big)$$

where the variance of $\hat{x}_{0l}$ is approximated using the delta method. This yields

$$\mathrm{V}(\hat{x}_{0l}) \doteq \underline{x}_{0l}'^{T} I^{-1}(\underline{\theta}) \underline{x}_{0l}'$$

where $\underline{x}_{0l}'$ is the vector of derivatives of x_{0l} with respect to $\underline{\theta}$. Then $I_0^{-1}(\hat{\underline{\theta}})$ and $\underline{x}_{0l}'(\hat{\underline{\theta}})$ are used in place of $I^{-1}(\underline{\theta})$ and $\underline{x}_{0l}'(\underline{\theta})$, and confidence intervals for x_{0l} at a given time within a particular replication can be obtained. Method 2 (Fieller intervals) assumes that

$$[\hat{a}_l - x_{0l}\hat{\beta}_l] \sim N\big(0, \mathrm{V}(\hat{a}_l - x_{0l}\hat{\beta}_l)\big),$$

where $\hat{a}_l = \big[\log\log(2) - \hat{\gamma}_l \log(t) - \hat{\alpha}_l\big]$. Forming a probability statement about $\hat{a}_l - x_{0l}\hat{\beta}_l$ and solving the resulting quadratic in x_{0l} yields a confidence interval for x_{0l}. These confidence intervals for $x_{0l}(t)$, the log ED50, obtained by either method can then be transformed to obtain confidence intervals for the ED50 at a given t for a particular l.

4. Goodness-Of-Fit and Model Simplification

The observed data has a multinomial likelihood. Thus, under the model assumptions, expected cell counts estimated via maximum likelihood lead directly to goodness-of-fit tests. Letting $\hat{u}_{ijl} = n_{il}[F(t_j \mid x_i, \hat{\underline{\theta}}_l) - F(t_{j-1} \mid x_i, \hat{\underline{\theta}}_l)]$, providing n is large, we have

$$X^2 = \sum_{i,j,l} \frac{(u_{ijl} - \hat{u}_{ijl})^2}{\hat{u}_{ijl}} \doteq \chi^2_{L(km-3)},$$

and

$$G^2 = 2\sum_{i,j,l} u_{ijl}\log(\frac{u_{ijl}}{\hat{u}_{ijl}}) \doteq \chi^2_{L(km-3)},$$

where $i = 1,...,k$, $j = 1,...,m+1$, $l = 1,...,L$. Under the model assumptions the limiting chi-square distributions for both X^2 and G^2 will be good approximations provided not too many of the expected cell counts are small.

The G^2 statistic also gives a method for testing whether eliminating parameters, or representing a group of parameters by a single one, significantly affects the fit of the model. An example would be to compare the model with parameters $\underline{\theta}_1 = (\alpha_l, \beta_l, \gamma_l; l = 1,...,L)^T$ to the reduced model with parameters $\underline{\theta}_2 = (\alpha_l, \beta_l, \gamma; l = 1,...,L)^T$, thus $\gamma_l = \gamma$ for $(l = 1,...,L)$. In general, if Model 2 is a reduced version of Model 1, then

$$\Delta G^2 = G_2^2 - G_1^2 \doteq \chi^2_{[\text{dfModel2}-\text{dfModel1}]}.$$

In this example:

$$\text{dfModel2} - \text{dfModel1} = [Lkm - 2L - 1] - [Lkm - 3L] = L - 1.$$

This yields a method for determining the most parsimonious model.

5. Application To An Example

In the previously described example presented in Carter and Hubert[3], each of $k = 7$ different concentrations of a toxic copper substance was administered to a single tank of $n_{il} = 10$ fish. The cumulative mortality counts for each tank were recorded on each of $m = 3$ days ($t_0 = 0$, $t_1 = 2$, $t_2 = 3$,

$t_3 = 4$) and the experiment had $L = 2$ replications. Carter and Hubert fit their model with $q = 2$ (linear in inverse time). The resulting model,

$$\hat{Z}_{ijl} = 1.193 - 1.286t_j^{-1} + 0.782x_i - 0.935x_it_j^{-1} + \hat{\rho}_{1l} + \hat{\rho}_{2l}t_j^{-1}$$

with $\hat{\rho}_{11} = -\hat{\rho}_{12} = 0.0302$ and $\hat{\rho}_{21} = -\hat{\rho}_{22} = -0.0853$, has 6 parameters, 2 of which model replication effects. Table VII gives the cumulative counts reported in the experiment, and the resulting estimates using the above model (shown in parentheses).

The proposed model was fit to this data with $\underline{\theta}_1 = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2)$ and seemed to give a very good fit ($X^2 = 14.49$, $G^2 = 18.01$, $df = 36$). Nested models, with parameters combined as in the previous example, were tested to see if the combining of the parameters significantly affected the fit, using G^2. It seemed reasonable to reduce the model first to $\underline{\theta}_2 = (\alpha_1, \beta_1, \alpha_2, \beta_2, \gamma)$ by combining γ_1 and γ_2 into one parameter. If this reduction did not significantly affect the fit, it would indicate that the relationship between time and the probability of response is the same on the 2 replications. Further reducing the model to $\underline{\theta}_3 = (\alpha_1, \alpha_2, \beta, \gamma)$ would similarily examine the relationship between dose and the probability of response over the 2 replications. Finally reducing the model to $\underline{\theta}_4 = (\alpha, \beta, \gamma)$ would indicate whether there are any replication effects. The parameters being estimated, the X^2 and G^2 values for each model, and the difference of the G^2's between nested models, for all of the models described above, are summa-

rized in Table VIII. Table VIII also examines the model reductions obtained when the γ parameter is fixed at $\gamma = 1$, both at the $\underline{\theta}_2 = (\alpha_1, \beta_1, \alpha_2, \beta_2, \gamma)$ stage and at the $\underline{\theta}_4 = (\alpha, \beta, \gamma)$ stage, to reduce the model from Weibull to exponential. The most parsimonious model was found to be the 3 parameter model with $\underline{\theta}_4 = (\alpha, \beta, \gamma)$, which suggests no replication effect. Table VIII also shows the significance of the γ parameter, and thus the Weibull generalization.*

The fitted cumulative mortality counts under Model 4 are given by

$$\hat{R}_{ijl} = n_{il}\left[1 - \exp(-\hat{\lambda}(x_i)t_j^{\hat{\gamma}})\right]$$

with $\hat{\lambda}(x_i) = \exp(\hat{\alpha} + \hat{\beta}x_i)$, $n_{il} = 10$, $\hat{\alpha} = -2.398$, $\hat{\beta} = 0.889$, and $\hat{\gamma} = 1.630$. The estimated covariance matrix for $\underline{\hat{\theta}} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma})$ is

$$I_0^{-1}(\underline{\hat{\theta}}) = \begin{pmatrix} 0.1103 & -0.0082 & -0.0764 \\ -0.0082 & 0.0188 & 0.0053 \\ -0.0764 & 0.0053 & 0.0620 \end{pmatrix}.$$

Table IX gives the observed cumulative counts and the cumulative counts estimated using Model 4 (shown in parentheses) which agree closely with the observed counts. Figure 3 shows the fitted dose response curves under Model 4 at various fixed time points, and illustrates the general shape of the fitted model. Figures 4, 5, and 6 show the fitted dose response curves

* It should be noted that, due to the small observed cell counts, the p-values should not be considered as exact probabilities; nevertheless, the summary provided in Table VIII is very clear.

at $t = 2.0$, 3.0, and 4.0 for both Carter and Hubert's model and Model 4, and illustrate the difference in shape of the two models. Comparing the fits as summarized in Tables 7 and 9 is difficult, but neither model seems to clearly fit the data better. One method for comparing the fit of the 2 models is to look at X^2 and G^2 calculated from the expected cell probabilities for each model. Though X^2 and G^2 for Carter and Hubert's model cannot be compared to a chi-squared distribution, they can be used as a crude measure of the extent of departure from the observed data. For Carter and Hubert's model $X^2 = 17.21$ and $G^2 = 18.99$, while for Model 4 the analogous values are $X^2 = 14.76$ and $G^2 = 18.42$. Not only does Model 4 yield smaller X^2 and G^2 values, but the model has only 3 parameters as compared to Carter and Hubert's 6 parameters; overall Model 4 seems to fit the data somewhat better than Carter and Hubert's model.

The ultimate objective of these analyses is inference for the ED50. Table X gives point estimates and 95% confidence intervals for the ED50 at some specific time points under Carter and Hubert's fitted model, and our fitted Model 4; the intervals are obtained using Fieller's method. Figure 7 shows pointwise confidence bands for the log ED50 over time for Model 4 using the two methods mentioned; the delta method and Fieller's method. Figure 8 shows confidence bands for the ED50 over time for both models, using Fieller's method. It can be seen that the confidence intervals obtained using

Model 4 are substantially wider than those obtained by Carter and Hubert. Model 4 seems a reasonable model and fits the data very well. There is no reason to believe that the confidence intervals obtained are grossly incorrect in their coverage probabilities, yet if these coverage probabilities are accurate then Carter and Hubert's model is badly overestimating the accuracy of its estimate of the ED50 at a fixed time point.

In general it would not seem advisable to use the model proposed by Carter and Hubert[3] in the example they presented. There are some major doubts as to the adequacy of the trivariate normal approximation being used in this situation. Also the confidence intervals obtained using their model in this example appear to be misleading. The approach illustrated in this chapter is a simpler and more direct approach to a problem of this type.

Table I
3pt Designs-Symmetric About ED50

$P(x) = 0.05$		
λ	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.0	4.000	1.00
0.1	4.773	0.84
0.2	5.917	0.68
0.3	7.781	0.51
0.4	11.362	0.35
0.5	21.044	0.19

$P(x) = 0.1$		
λ	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.0	4.000	1.00
0.1	4.587	0.87
0.2	5.376	0.74
0.3	6.493	0.62
0.4	8.196	0.48
0.5	11.109	0.36

$P(x) = 0.2$		
λ	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.0	4.000	1.00
0.1	4.310	0.93
0.2	4.673	0.86
0.3	5.102	0.78
0.4	5.617	0.71
0.5	6.249	0.64

$P(x) = 0.3$		
λ	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.0	4.000	1.00
0.1	4.132	0.97
0.2	4.273	0.93
0.3	4.424	0.90
0.4	4.587	0.87
0.5	4.761	0.84

Table II
5pt Designs- Symmetric About ED50

$P(x_1) = 0.1\ P(x_2) = 0.2$			
λ_1	λ_2	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.05	0.20	5.05	0.79
0.10	0.20	5.49	0.73
0.15	0.20	6.02	0.66
0.20	0.20	6.67	0.60

$P(x_1) = 0.1\ P(x_2) = 0.3$			
λ_1	λ_2	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.05	0.20	4.59	0.87
0.10	0.20	4.95	0.81
0.15	0.20	5.38	0.74
0.20	0.20	5.88	0.68

$P(x_1) = 0.2\ P(x_2) = 0.3$			
λ_1	λ_2	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.05	0.20	4.44	0.90
0.10	0.20	4.63	0.86
0.15	0.20	4.83	0.83
0.20	0.20	5.05	0.79

$P(x_1) = 0.2\ P(x_2) = 0.4$			
λ_1	λ_2	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.05	0.20	4.22	0.95
0.10	0.20	4.39	0.91
0.15	0.20	4.57	0.88
0.20	0.20	4.76	0.84

Table III
Spt Design $P(x) = 0.2$
Incorrect Value Of ED50

$P(x_0) = 0.3$		
λ	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.1	14.37	0.28
0.2	8.79	0.46
0.3	7.15	0.56
0.4	6.55	0.61

$P(x_0) = 0.4$		
λ	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.1	6.56	0.61
0.2	5.57	0.72
0.3	5.53	0.72
0.4	5.79	0.69

$P(x_0) = 0.42$		
λ	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.1	5.74	0.70
0.2	5.24	0.76
0.3	5.37	0.75
0.4	5.72	0.70

$P(x_0) = 0.44$		
λ	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.1	5.11	0.78
0.2	4.99	0.80
0.3	5.25	0.76
0.4	5.68	0.70

$P(x_0) = 0.46$		
λ	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.1	4.66	0.86
0.2	4.81	0.83
0.3	5.17	0.77
0.4	5.64	0.71

Table IV
5pt Designs $P(x_1) = 0.2$ $P(x_2) = 0.4$
Incorrect Value Of ED50

$P(x_0) = 0.3$			
λ_1	λ_2	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.05	0.20	18.73	0.21
0.10	0.20	12.25	0.33
0.15	0.20	9.61	0.41
0.20	0.20	8.23	0.49

$P(x_0) = 0.4$			
λ_1	λ_2	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.05	0.20	7.36	0.54
0.10	0.20	6.09	0.66
0.15	0.20	5.65	0.71
0.20	0.20	5.50	0.73

$P(x_0) = 0.42$			
λ_1	λ_2	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.05	0.20	6.19	0.65
0.10	0.20	5.46	0.73
0.15	0.20	5.25	0.76
0.20	0.20	5.22	0.77

$P(x_0) = 0.44$			
λ_1	λ_2	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.05	0.20	5.32	0.75
0.10	0.20	4.98	0.80
0.15	0.20	4.95	0.81
0.20	0.20	5.02	0.80

$P(x_0) = 0.46$			
λ_1	λ_2	$(n/\sigma^2)V(\hat{\mu})$	Efficiency
0.05	0.20	4.70	0.85
0.10	0.20	4.65	0.86
0.15	0.20	4.73	0.85
0.20	0.20	4.87	0.82

Table V
Regional Comparison Of Study Design To 5pt Design

Question	Region	$\hat{\mu}$	$\hat{\sigma}$	$V(\hat{\mu})$-study	$V(\hat{\mu})$-5pt
2	Victoria	60.50	-30.72	45.43	5.00
	P Alberni	138.14	-59.06	49.56	8.38
	Campbell R	79.72	-68.89	94.87	37.72
	Sechelt	60.94	-26.85	13.47	8.62
4	Victoria	21.11	-8.19	0.59	0.34
	P Alberni	73.18	-24.07	8.18	1.25
	Campbell R	35.60	-13.63	3.24	1.21
	Sechelt	41.34	-13.05	7.40	1.88

Table VI
Comparison Of 5pt Equal Allocation Design to Study Design For Various Values Of μ

μ	$V(\hat{\mu}) - 5pt$	$V(\hat{\mu})$-study design
10	8.02	8.12
20	4.92	4.85
30	4.01	3.86
40	3.80	3.68
50	4.01	4.05
60	4.92	5.32
70	8.02	9.23

Table VII
Observed And Expected Cumulative
Mortality Counts; Carter and Hubert

Block	Time	Concentration($\mu g/l$)						
#	(*days*)	0.10	0.20	0.30	0.50	1.00	2.00	2.50
1	2	0(0.49)	1(0.98)	1(1.33)	2(1.84)	3(2.62)	3(3.49)	4(3.79)
	3	1(0.85)	1(1.79)	2(2.47)	3(3.42)	5(4.81)	6(6.21)	7(6.65)
	4	1(1.06)	1(2.28)	3(3.13)	4(4.30)	8(5.94)	7(7.48)	9(7.93)
2	2	0(0.60)	1(1.13)	1(1.50)	2(2.03)	3(2.84)	4(3.73)	4(4.03)
	3	1(0.83)	1(1.77)	2(2.44)	3(3.39)	5(4.77)	6(6.18)	7(6.61)
	4	1(0.96)	1(2.13)	2(2.97)	4(4.13)	7(5.77)	8(7.33)	8(7.79)

Table VIII
Testing Nested Models

Model #	Parameters	X^2	G^2	*df*	*p*	ΔG^2	*df*	*p*
1	$(\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2)$	14.49	18.01	36	*large*	—	-	—
2	$(\alpha_1, \beta_1, \alpha_2, \beta_2, \gamma)$	14.77	18.33	37	*large*	0.32	1	>0.1
2*	$(\alpha_1, \beta_1, \alpha_2, \beta_2, 1)$	27.25	26.72	38	*large*	8.39	1	<0.01
3	$(\alpha_1, \alpha_2, \beta, \gamma)$	14.75	18.37	38	*large*	0.04	1	>0.1
4	(α, β, γ)	14.76	18.42	39	*large*	0.05	1	>0.1
4*	$(\alpha, \beta, 1)$	27.29	26.81	40	*large*	8.39	1	<0.01

Table IX
Observed And Expected Cumulative
Mortality Counts; Model 4

Block	Time	Concentration($\mu g/l$)						
#	(*days*)	0.10	0.20	0.30	0.50	1.00	2.00	2.50
1	2	0(0.36)	1(0.65)	1(0.92)	2(1.41)	3(2.45)	3(4.06)	4(4.70)
	3	1(0.68)	1(1.22)	2(1.70)	3(2.55)	5(4.20)	6(6.35)	7(7.08)
	4	1(1.06)	1(1.88)	3(2.58)	4(3.75)	8(5.81)	7(8.01)	9(8.60)
2	2	0(0.36)	1(0.65)	1(0.92)	2(1.41)	3(2.45)	4(4.06)	4(4.70)
	3	1(0.68)	1(1.22)	2(1.70)	3(2.55)	5(4.20)	6(6.35)	7(7.08)
	4	1(1.06)	1(1.88)	2(2.58)	4(3.75)	7(5.81)	8(8.01)	8(8.60)

Table X
Point Estimates And Confidence Intervals For ED50

Time	$\hat{x}_0$(Model 4)	95% CI	$\hat{x}_0$(C & H)	95% CI
2.00	2.76	(1.82,4.88)	5.60	(3.99,8.76)
2.50	1.83	(1.32,2.77)	1.83	(1.61,2.11)
3.00	1.31	(0.98,1.81)	1.02	(0.93,1.38)
3.50	0.99	(0.73,1.32)	0.84	(0.68,1.07)
4.00	0.77	(0.54,1.04)	0.70	(0.56,0.91)
6.00	0.37	(0.20,0.57)	0.49	(0.37,0.66)

Figure 1

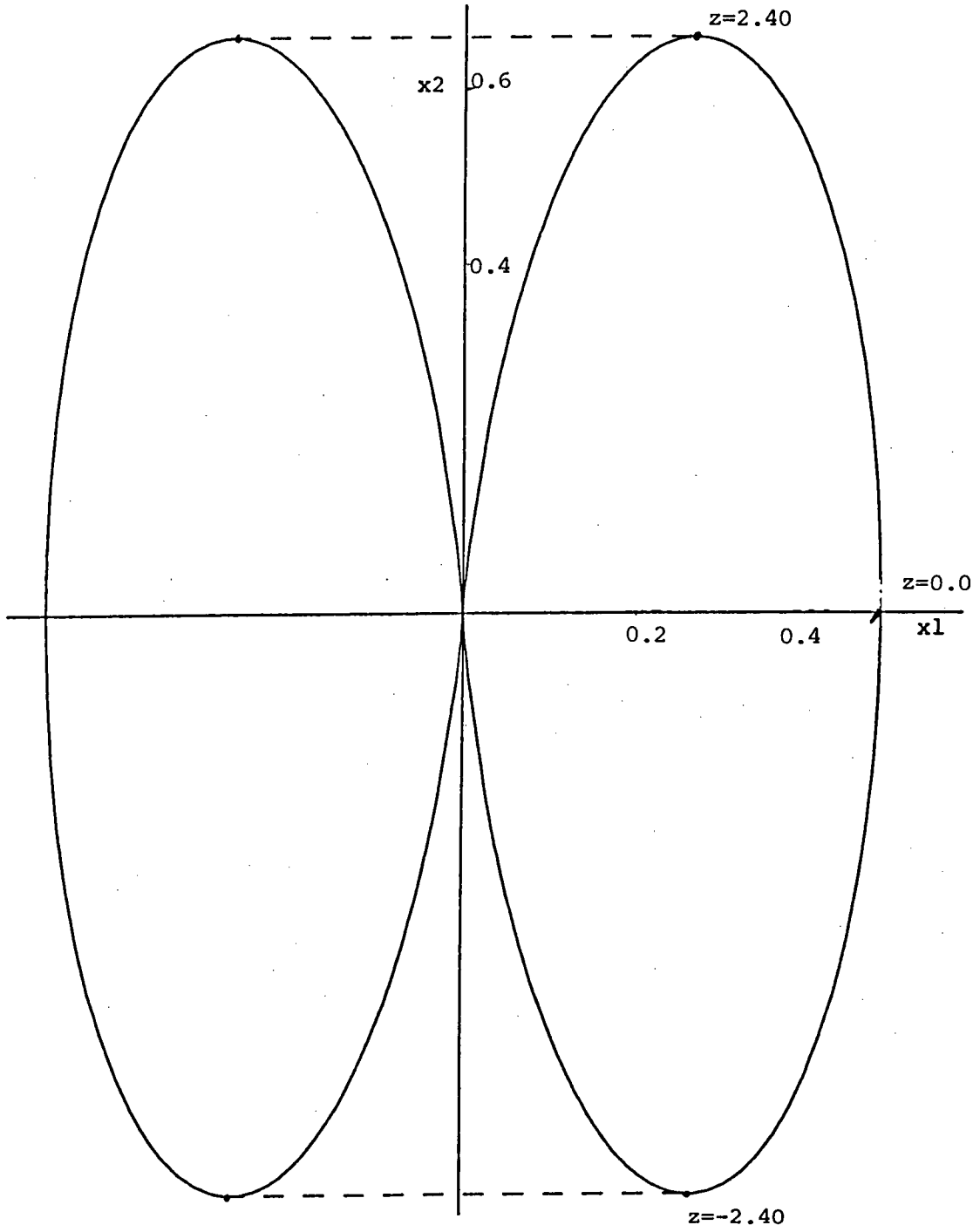


Figure 2

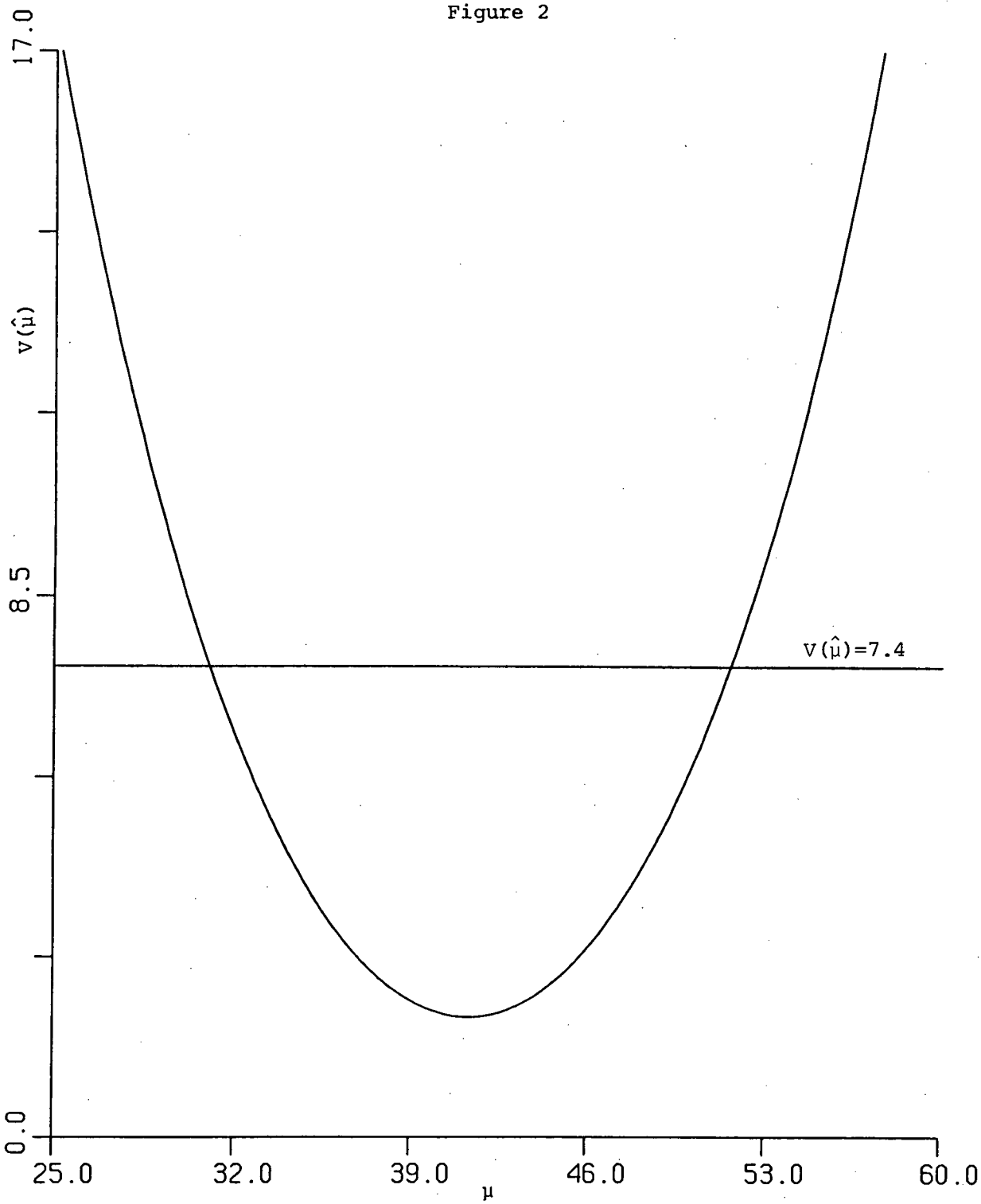


Figure 3

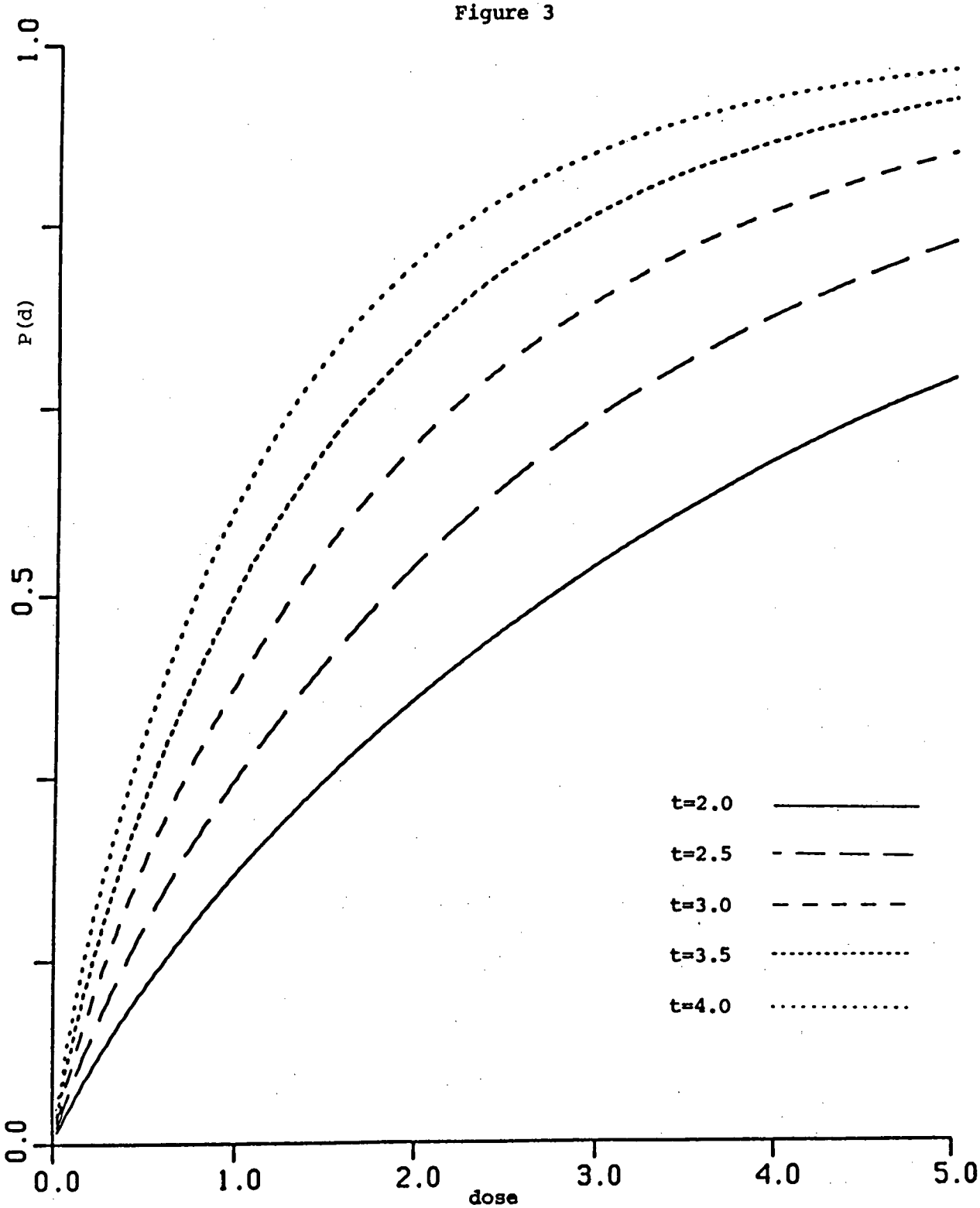


Figure 4

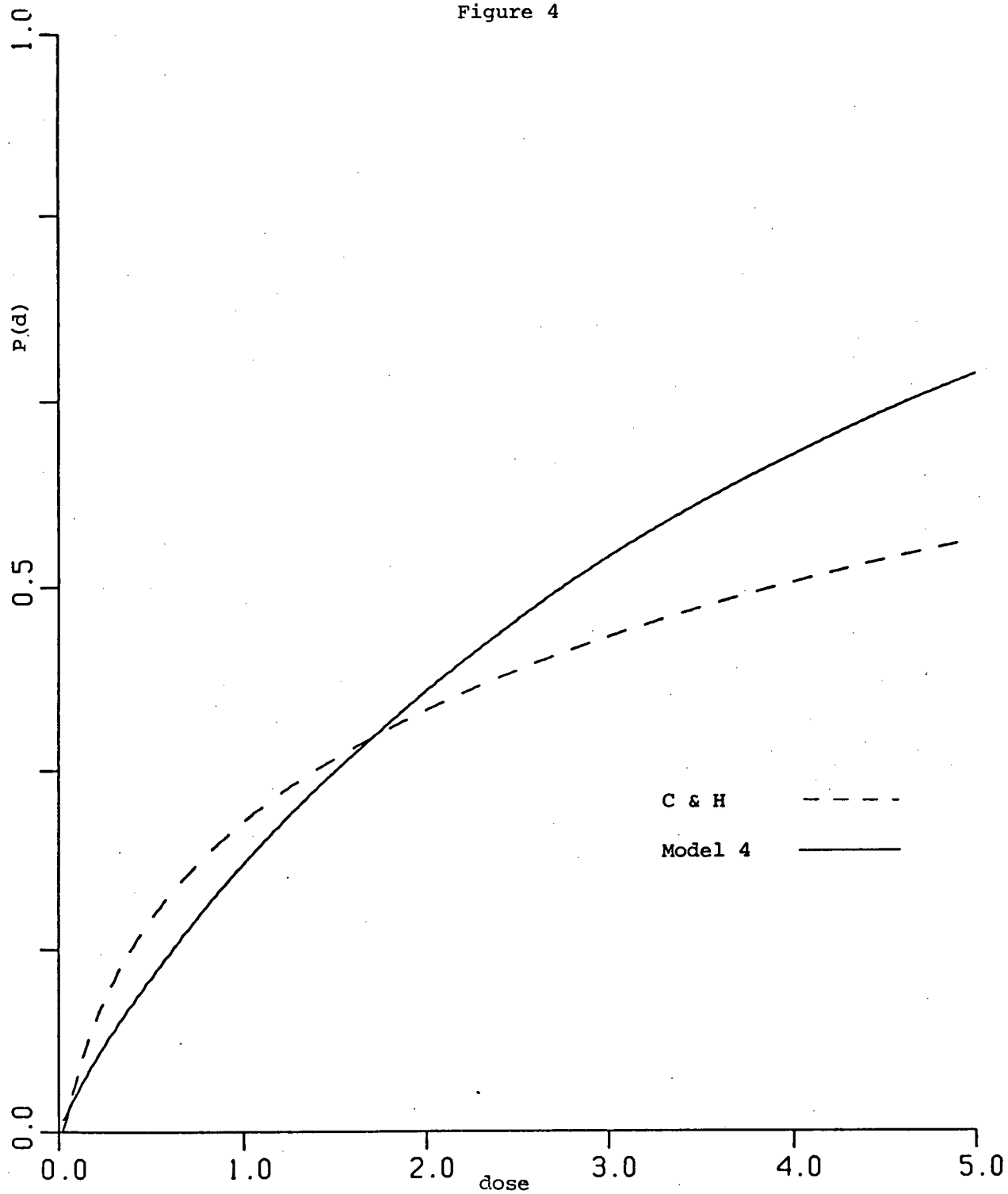


Figure 5

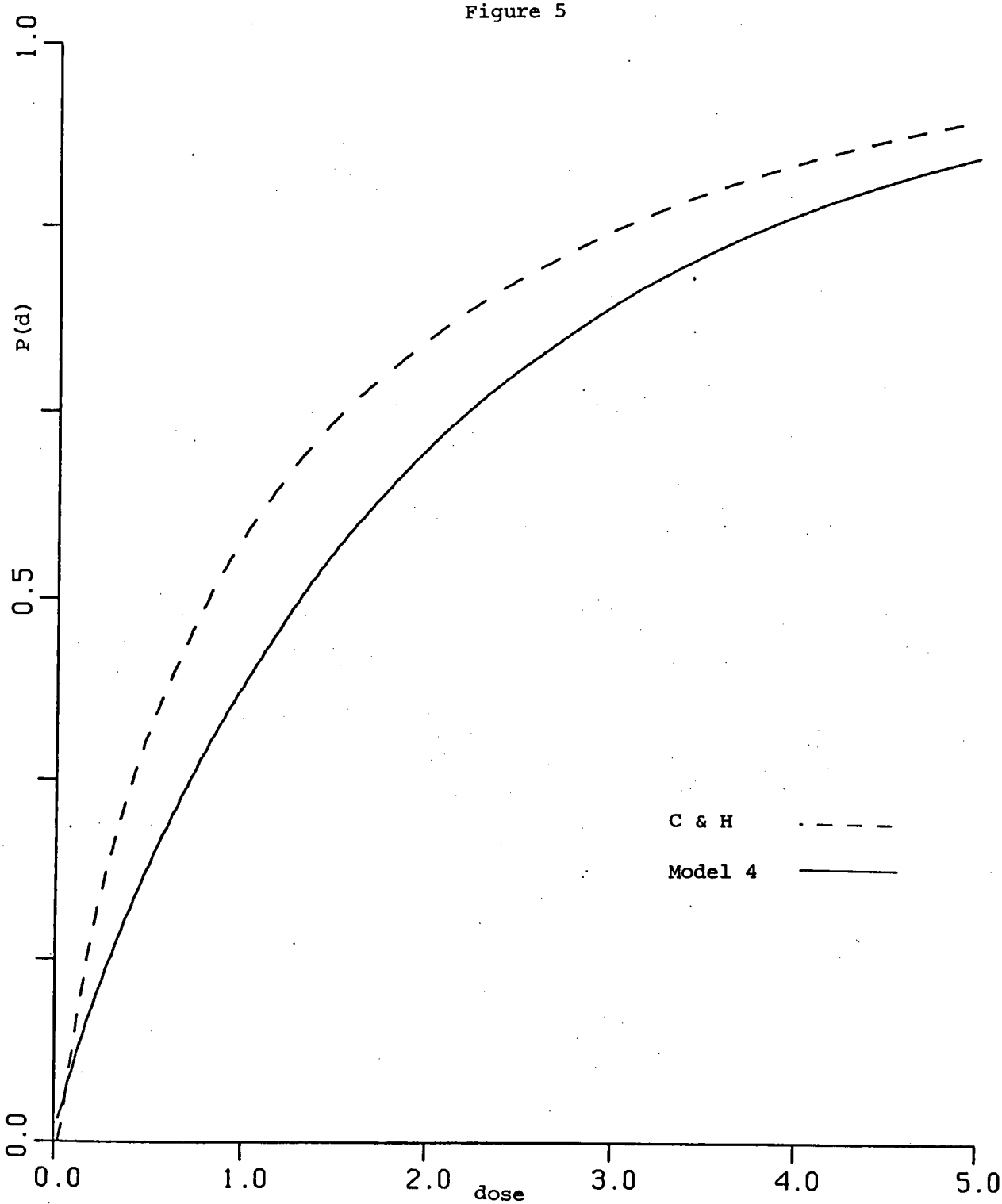


Figure 6

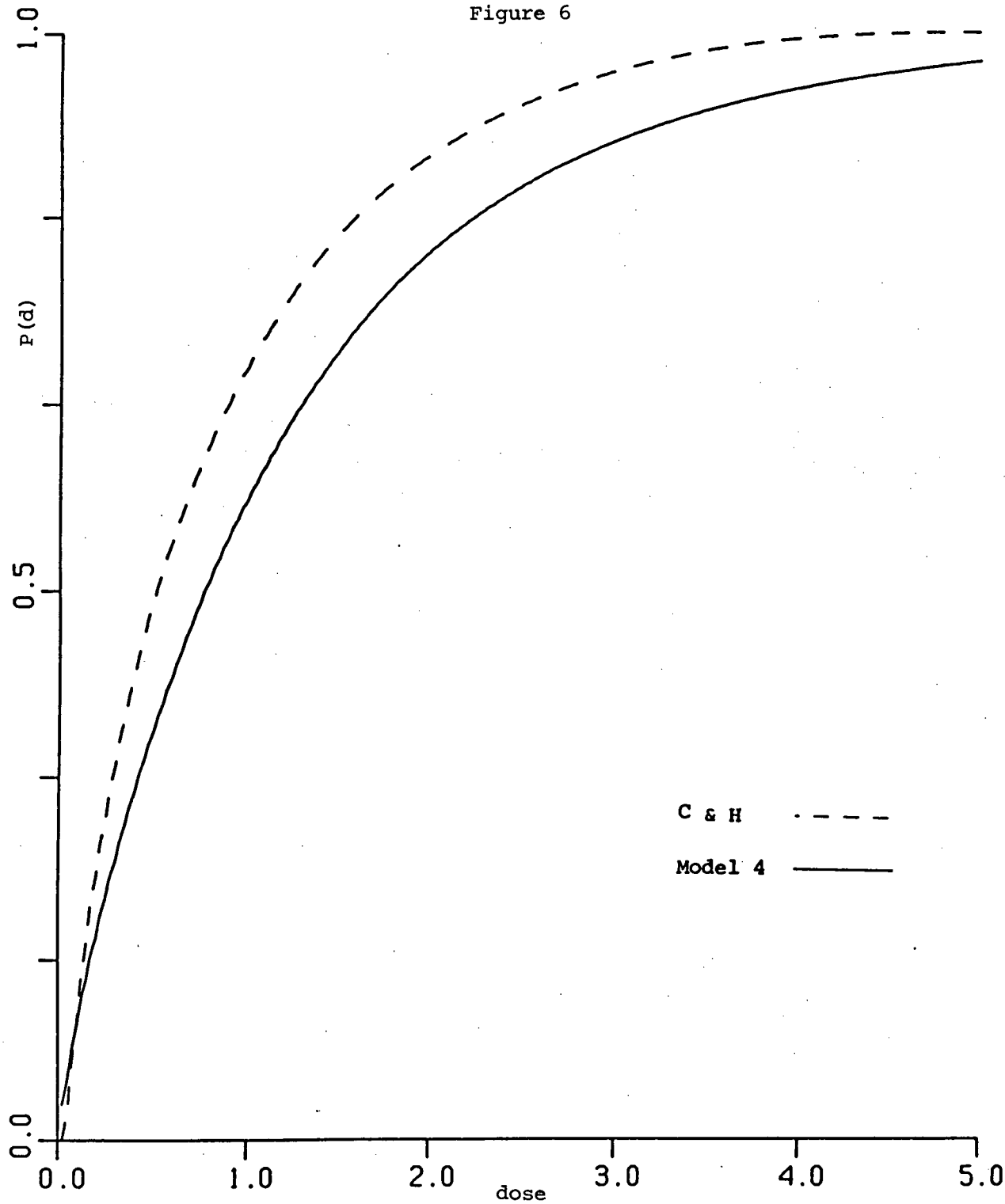


Figure 7

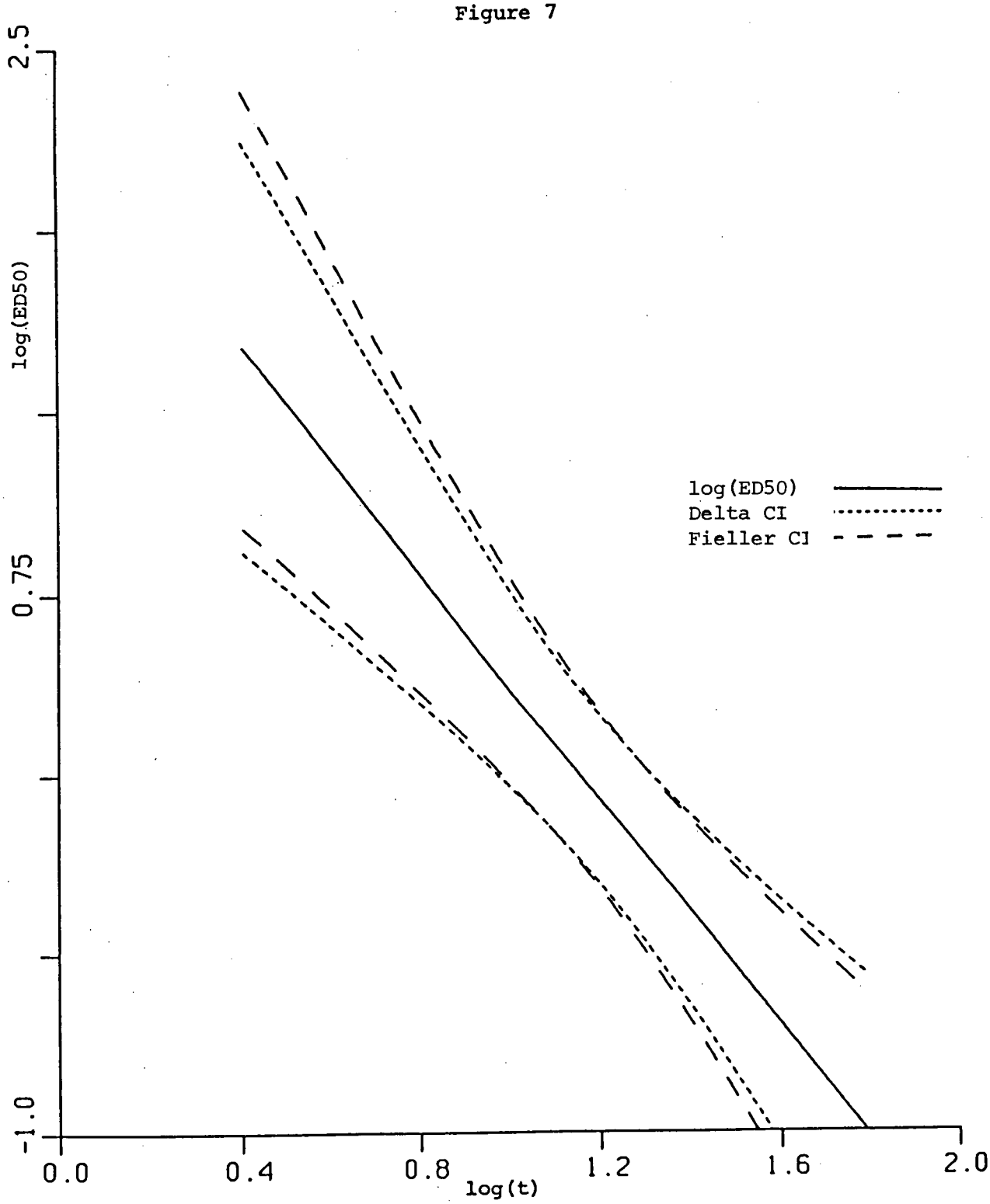
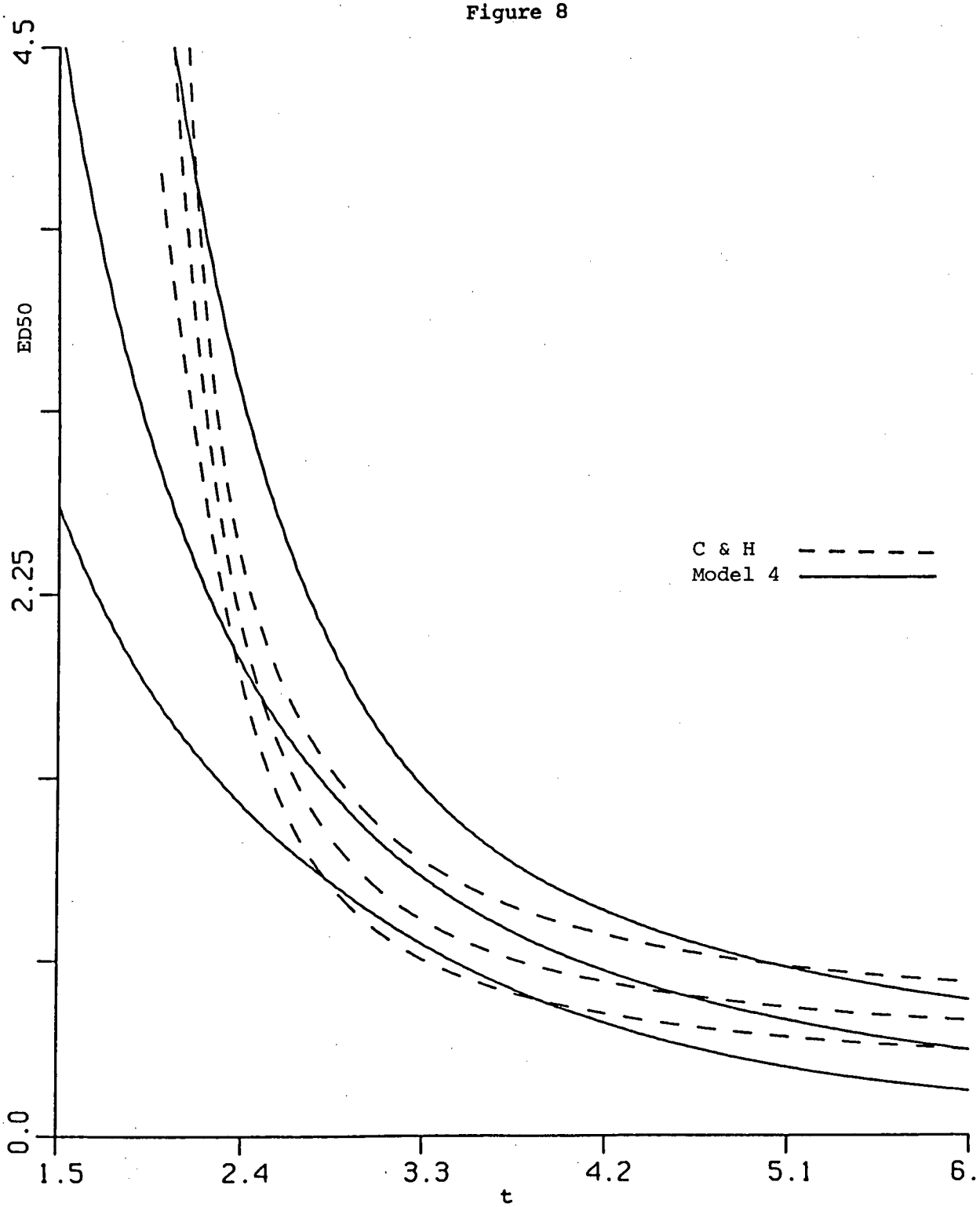

2.5
0.75
-1.0
log(ED50)
0.0
0.4
0.8
1.2
1.6
2.0
log(t)
log(ED50)
Delta CI
Fieller CI

Figure 8



References

[1] Berkson, J. (1953). "A Statistically Precise and Relatively Simple Method of Estimating the Bioassay With Quantal Response,Based on the Logistic Function". Journal of American Statistical Association, 48, 565–599.

[2] Brier, S. S. (1980). "Analysis of Contingency Tables Under Cluster Sampling". Biometrika, 67, 591–596.

[3] Carter, E. M. & Hubert, J. J. (1984). "A Growth-Curve Model Approach to Multivariate Quantal Bioassay". Biometrics, 40, 699–706.

[4] Chapman, D. G. & Nam, J. (1968). "Asymptotic Power of Chi Square Tests for Linear Trends in Proportions". Biometrics, 24, 315–328. Corrections- Biometrics, 25, 777.

[5] Chernoff, H. (1972). "Sequential Analysis and Optimal Designs". Philadelphia: Society for Industrial and Applied Mathematics.

[6] Elfving, G. (1952). "Optimum Allocation in Linear Regression Theory". Annals of Mathematical Statistics, 23, 255–262.

[7] Finney, D. J. (1971). "Probit Analysis(3rd ed.). Cambridge: Cambridge University Press.

[8] Kalish, L. A. & Rosenberger, J. L. (1978). "Optimal Designs for the Estimation of the Logistic Function". Technical Report 33, Department of Statistics, Pennsylvania State University.

[9] Kuo, L. (1983). "Bayesian Bioassay Design". Annals of Statistics, 11, 886–895.

[10] Mitra, S. K. (1958). "On The Limiting Power Function of the Frequency χ^2 Test". Annals of Mathematical Statistics, 29, 1221–1233.

[11] Moismann, J. E. (1962). "On the Compound Multinomial Distribution, the Multivariate β-distribution, and Correlation Among Proportions." Biometrika, 49, 65–82.

[12] Potthoff, R. F. & Roy, S. N. (1964). "A Generalized Multivariate Analysis of Variance Model Useful Especially for Growth Curve Problems". Biometrika, 51, 313–326.

[13] Smith, K. C., Savin, N. E., & Robertson, J. L. (1984). "A Monte Carlo Comparison of Maximum Likelihood and Minimum Chi Square Sampling Distributions in Logit Analysis". Biometrics, 40, 471–482.

[14] Wetherill, G. B. (1962). "Sequential Estimation of Quantal Response Curves". Journal of the Royal Statistical Society, Ser. B, 25, 1–48.

[15] Wilson, J. R. (1986). " Approximate Distribution and Test of Fit for the Clustering Effect in the Dirichlet Multinomial Model". Communications in Statistics: Theory and Methods, 15, 1235–1249.