New Developments and Applications in Experimental Design
IMS Lecture Notes – Monograph Series (1998) Volume 34

D-OPTIMAL DESIGNS GIVEN A BIVARIATE PROBIT RESPONSE FUNCTION

BY ELOI P. KPAMEGAN[1]

American University

We consider the design of an experiment involving two drugs; associated with each drug is an identifiable type of toxicity. For example, in cancer chemotherapy, cyclophosphamide is toxic to the heart and busulfan is toxic to the liver. Therefore, four responses are possible: no toxicity, toxicity to the heart, toxicity to the liver and toxicity to both organs. Consequently, responses are bivariate binary random variables that depend on a bivariate treatment space. Assuming a bivariate probit response function with $0 \leq \rho < 1$, we find D-optimal designs when the responses are independent and give preliminary results for when they are dependent.

1. Introduction. An understanding of the action of mixtures of drugs is becoming increasingly important in a variety of scientific disciplines, ranging from pharmacology and toxicology on the one hand to industrial hygiene and environmental protection on the other. The action of even a single drug upon a biological organism involves, in general, a complex sequence of processes, and if more than one drug is present, the situation is further complicated. The role of mathematical models in this context is now recognized; see, for example, Hewlett and Plackett (1959, 1964, 1979); Ashford and Smith (1964, 1965); Ashford and Cobby (1974). A *response function* describes the relationship between response probabilities and explanatory or control variables.

Our work is motivated by Flournoy (1993). Her paper describes a pilot clinical trial that aims to find dosages of a new drug combination that will produce 10 percent toxicity, which was the toxicity rate produced by the standard single drug treatment. Two drugs were involved in the experiment: cyclophosphamide and busulfan. The drug combinations to be examined were restricted to lie on the line defined by the points $(x_1, x_2) = (40, 6)$ and $(x_1, x_2) = (180, 20)$, where x_1 denotes the dose of cyclophosphamide (mg/kg/day) and x_2 the dose of busulfan (mg/kg/day). Possible dosages of

---

[1] Research supported, in part, under cooperative agreement CR 822636-01 from the U.S. Environmental Protection Agency.

Received September 1997; revised August 1998.

AMS 1991 subject classifications. Primary 62K05.

Key words and phrases. Binary random variables, combination therapy, dose finding, experimental design, mixture of drugs.

cyclophosphamide, namely, 31, 49, 67, 85, 103 and 121, were identified according to the method of Tsutakawa (1980) which minimizes the posterior variance around the target percentile.

Empirical results from the experiment that used these dosages indicate that the dose-response function increases much more sharply than does a prior upper bound that was established by experts. The synergy between the drugs was much stronger than expected. A more informative experiment would have taken the possible interactions of the two drugs into account (a sequential approach would also have helped).

In this paper we explore optimal design theory for bivariate probit response functions, and thereby introduce formal design considerations for allocating subjects to treatments when there are two binary responses and possible drug interactions.

2. Optimal design for nonlinear problems. The design of experiments is an important part of scientific research. Design involves specifying how an experiment will be conducted and choosing the values of variables that can be controlled before the experiment starts. Although many processes are relevant aspects in the design of an experiment, we focus on choosing which treatments to study and the proportion of observations to allocate to each treatment.

Let ξ be the measure that determines the treatment allocation distribution across the design space. For example, consider a design variable x in R^1 and two specific values of x, namely x_1 and x_2. If half of the subjects are treated at x_1 and half at x_2, the measure ξ is zero almost everywhere, with spikes at x_1 and x_2, each of height $\frac{1}{2}$. We say that a point x is a *design point* if it has positive probability with respect to ξ. Of course, $\int \xi\,(dx) = 1$. In the example, the integration ranges over R^1. However, for designs developed in this paper, x will be a bivariate point and integration will range over R^2. Since ξ determines both which treatments are to be used and the proportion of subjects to be allocated to each treatment, we refer to ξ as the *design*.

The theory of optimal experimental design has been extensively studied for linear regression models by Silvey (1980), Fedorov (1972, 1981), and others. With nonlinear response functions, the optimal design depends on the unknown parameters. One approach to this problem is to design an experiment to be efficient for a best guess of the parameter values. This approach leads to "locally optimal" designs, a term introduced by Chernoff (1953). Bayesian optimal designs are a natural generalization of the locally optimal designs in which the criterion for selecting points is averaged with respect to a prior distribution on the unknown parameters rather than evaluated at a single guess [see review papers by Cochran (1973), Ford, Titterington, and Kitsos (1989) and Chaloner and Verdinelli (1995)]. An excellent overall review of c-optimal and D-optimal experimental designs is given by Atkinson (1996).

We focus on the D-optimality criterion which is useful when one wishes to estimate all parameters in the response surface model and considers each to be of equal significance. Let $L\,(\theta|\xi)$ denote the likelihood given an experiment with design ξ where $\theta = (\theta_{\mathbf{1}}, \ldots, \theta_{\mathbf{p}})^{\mathbf{T}}$ are the parameters of the response surface. Let $\mathbf{I}(\xi;\theta)$ be the $p \times p$ dimensional Fisher information matrix with elements

$$I_{sk}(\xi;\theta) = -\mathbf{E}\left[\mathbf{d^2} \log \mathbf{L}\,(\theta|\xi) \div \mathbf{d}\theta_{\mathbf{s}}\mathbf{d}\theta_{\mathbf{k}}\right],$$

assuming that the derivatives exist, and let $\det[\mathbf{I}(\xi;\theta)]$ denote the determinant of the Fisher information matrix. Then ξ^D is said to be D-optimal if it maximizes $\det[\mathbf{I}(\xi;\theta)]$ with respect to ξ.

Little work has been done on optimal designs for binary response experiments involving more than one dimension for explanatory or response variables. To the best of our knowledge, no previous work exists where both explanatory and responses variables are multivariate. Heise and Myers (1996) extend the general equivalence theorem to accommodate multivariate response variables. Sitter and Torsney (1995) show there exists a D-optimal design for binary response experiments with two design variables where the probability of response is modeled by a nonlinear location location-scale distribution function. Atkinson (1996) finds locally optimal and Bayesian designs for estimating the 95th percentile of a logistic function which models the probability of death in an experiment on flies. Here we consider bivariate binary response experiments with two design variables.

Although optimal designs for nonlinear models are typically impossible to achieve explicitly in practice, we are still interested in constructing them because they provide a reference point for evaluating the performance of other designs. Furthermore, sequential allocation strategies can be devised that cluster around or converge to the optimal design points. For example, let θ be the parameters of the response function. Then if a treatment can be allocated sequentially in batches, Hall (1968) suggests that the design points used for batch $\eta+1$ be elected according to an estimated optimal design based on $\widehat{\theta}_\eta$, where $\widehat{\theta}_\eta$ is the estimate of θ based on the first η batches. Alternatively, up and down procedures can be constructed so as to cluster the distribution of treatments around the optimal design points [Derman (1957), Durham and Flournoy (1994, 1995), Giovagnoli and Pintacuda (1996, 1998)].

3. Bivariate response functions with two stimuli. Let $(x_i\ ,\ x_j)$ denote a treatment consisting of the ith level of drug A and the jth level of drug B, $i=1,\ldots,r$; $j=1,\ldots,c$. Let Y_0 and Y_1 be Bernoulli random variables for responses of type 0 and type 1. In our example, $\mathrm{Y}_k=1$ if a type k toxic response occurs, and $\mathrm{Y}_k=0$ otherwise, $k=0,1$. In particular, the response Y_0 at $(x_i\ ,\ x_j)$ indicates whether or not the $(x_i\ ,\ x_j)$ combination of busulfan and cyclophosphamide results in toxicity to the heart and the response Y_1 at $(x_i\ ,\ x_j)$ indicates whether $(x_i\ ,\ x_j)$ results in toxicity to the liver.

Consider bivariate response functions that can be described by bivariate cumulative distribution functions. Writing $(i,j) := (x_i, x_j)$ and suppressing reference to dependencies on θ, the possible response probabilities at each stimuli (x_i, x_j) are

$$
\begin{aligned}
F_{11}(i,j) &: = P(Y_0=1, Y_1=1 \mid (x_i,x_j));\\
F_{10}(i,j) &: = P(Y_0=1, Y_1=0 \mid (x_i,x_j));\\
F_{01}(i,j) &: = P(Y_0=0, Y_1=1 \mid (x_i,x_j));\\
F_{00}(i,j) &: = P(Y_0=0, Y_1=0 \mid (x_i,x_j));\\
F_{00}(i,j) &= 1-F_{11}(i,j)-F_{10}(i,j)-F_{01}(i,j).
\end{aligned}
\tag{3.1}
$$

The likelihood function for a single trial at (x_i, x_j) is

$$L\left[\theta \mid (\mathbf{x_i}, \mathbf{x_j})\right] = \left[F_{11}(i,j)\right]^{Y_0 Y_1} \left[F_{10}(i,j)\right]^{Y_0(1-Y_1)} \left[F_{01}(i,j)\right]^{Y_1(1-Y_0)} \times \left[F_{00}(i,j)\right]^{(1-Y_0)(1-Y_1)},$$

and the (s,k)th element of the Fisher information matrix for a single trial at (x_i, x_j) is

$$I_{sk}\left[(x_i, x_j); \theta\right] = -E\left\{\partial^2 \log L\left[\theta \mid (\mathbf{x_i}, \mathbf{x_j})\right] \div \partial\theta_s \partial\theta_k\right\}. \tag{3.2}$$

If $N(1,1), \ldots, N(r,c)$ independent trials are performed at $(x_1, x_1), \ldots, (x_r, x_c)$, respectively, and $\sum_i \sum_j N(i,j) = n$, the (s,k)th element of the Fisher information matrix is

$$I_{sk}\left[\xi(\mathbf{x}); \theta\right] = \sum_i^r \sum_j^c N(i,j)\, I_{sk}\left[(x_i, x_j); \theta\right] = n \sum_i^r \sum_j^c \xi_{ij} I_{sk}\left[(x_i, x_j); \theta\right], \tag{3.3}$$

where $\xi_{ij} = N(i,j) \div n$.

4. Finding the *D*-optimal design in a multivariate setting. It is well established that if the design space spans R^p, and if a D-optimal design measure is supported by p points, then it puts probability $1 \div p$ at each of them (Silvey, 1980). This result holds in a multivariate setting. The general equivalence theorem [Kiefer and Wolfowitz, 1960] is useful for verifying that a design is D-optimal. It was recently extended to accommodate multivariate response variables by Heise and Myers (1996). The further extension to accommodate multivariate design variables as well is trivial and is given as follows for $\mathbf{X} \subset R^2$.

THEOREM 4.1 *The following three assertions are equivalent:*

1. ξ^D *maximizes* $\det\left[\mathbf{I}(\xi; \theta)\right]$ *over all* ξ *in* $\mathbf{X}$.
2. ξ^D *minimizes* $\max_{(x_i, x_j) \in \mathbf{R^2}} \left\{trace\left[\mathbf{I^{-1}}\left[\xi(\mathbf{x}); \theta\right] \mathbf{I}\left[\xi(\mathbf{x_i}, \mathbf{x_j}); \theta\right]\right]\right\}$ *over all* ξ *in* $\mathbf{X}$, *where the elements of* $\mathbf{I}\left[\xi(\mathbf{x_i}, \mathbf{x_j}); \theta\right]$ *are given by* (3.2) *and the elements of* $\mathbf{I}\left[\xi(\mathbf{x}); \theta\right]$ *are given by* (3.3).
3. $trace\left\{\mathbf{I^{-1}}\left[\xi^{\mathbf{D}}(\mathbf{x}); \theta\right] \mathbf{I}\left[\xi(\mathbf{x_i}, \mathbf{x_j}); \theta\right]\right\} \leq p$ *where* p *is the number of parameters in the model, and equality is achieved when* (x_i, x_j) *are the optimal design points.*

This theorem states that to verify that a particular design ξ is optimum it is sufficient to show that $\text{trace}\left\{\mathbf{I^{-1}}\left[\xi^{\mathbf{D}}(\mathbf{x}); \theta\right] \mathbf{I}\left[\xi(\mathbf{x_i}, \mathbf{x_j}); \theta\right]\right\}$ achieves a maximum value of p at the optimal design points. If the optimal design is found by numerical optimization, the equivalence theorem provides a method for verifying that the design is globally optimal.

For the bivariate probit response function, $F_{11}(i,j) = \int_{-\infty}^{x_i} \int_{-\infty}^{x_j} \phi(u,v) du dv$, the cumulative bivariate normal distribution function, where $\phi(u,v)$ is the bivariate normal probability density function. Specifically,

$$\phi(u,v) = \left(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}\right)^{-1} \exp\left\{-Q(u,v)/2(1-\rho^2)\right\},$$

$-\infty < u < \infty$, $-\infty < v < \infty$, $\sigma_1 > 0$, $\sigma_2 > 0$, $|\rho| < 1$, and

$$Q(u,v) = \left(\frac{u-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{u-\mu_1}{\sigma_1}\right)\left(\frac{v-\mu_2}{\sigma_2}\right) + \left(\frac{v-\mu_2}{\sigma_2}\right)^2.$$

We describe the D-optimal design for independent probit responses in Section 5 and D-optimal design for dependent probit responses in Section 6.

5. Independent probit drug action. Let $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-u^2/2)\, du$ be the cumulative normal distribution function. In the special case where two binary responses follow independent probit response functions, Fisher's information matrix is a block diagonal matrix which can be written in the form

$$\mathbf{I}(\xi;\theta) = \mathbf{n}\sum_{\mathbf{i=1}}^{\mathbf{r}}\sum_{\mathbf{j=1}}^{\mathbf{c}}\xi_{\mathbf{ij}} \begin{bmatrix} \frac{1}{\sigma_1^2}w(z_i) & \frac{1}{\sigma_1^2}z_i w(z_i) & 0 & 0 \\ \frac{1}{\sigma_1^2}z_i w(z_i) & \frac{1}{\sigma_1^2}z_i^2 w(z_i) & 0 & 0 \\ 0 & 0 & \frac{1}{\sigma_2^2}w(u_j) & \frac{1}{\sigma_2^2}u_j w(u_j) \\ 0 & 0 & \frac{1}{\sigma_2^2}u_j w(u_j) & \frac{1}{\sigma_2^2}u_j^2 w(u_j) \end{bmatrix} \tag{5.1}$$

where $z_i = (x_i - \mu_1) \div \sigma_1$, $u_j = (x_j - \mu_2) \div \sigma_2$, and

$$w(\cdot) = \phi^2(\cdot) \div \Phi(\cdot)[1 - \Phi(\cdot)].$$

Let $\mathbf{Z} \subset \mathbf{R}$ and $\mathbf{U} \subset \mathbf{R}$ be two standardized design spaces spanned by z_i and u_j, respectively, and let $\mathbf{\Omega_Z}$ and $\mathbf{\Omega_U}$ be their canonical design spaces [Ford, Torsney and Wu (1992)]. Let $\mathbf{z} \in \mathbf{Z^2}$ and $\mathbf{u} \in \mathbf{U^2}$. Denote by $\mathbf{I}[\xi(\mathbf{z})]$ and $\mathbf{I}[\xi(\mathbf{u})]$ Fisher's information matrices for the univariate probit responses. Then it follows from (5.1) that

$$\begin{aligned} \det\{\mathbf{I}[\xi;\theta]\} &= \frac{1}{\sigma_1^4\sigma_2^4}\det\{\mathbf{I}[\xi(\mathbf{z})]\} \times \det\{\mathbf{I}[\xi(\mathbf{u})]\} \\ &:= \frac{1}{\sigma_1^4\sigma_2^4}\det\{\mathbf{I}[\xi(\mathbf{z},\mathbf{u})]\}. \end{aligned} \tag{5.2}$$

The D-optimal design which maximizes $\det\{\mathbf{I}[\xi(\mathbf{z},\mathbf{u})]\} := \det\{\mathbf{I}[\xi(\mathbf{z})]\} \times \det\{\mathbf{I}[\xi(\mathbf{u})]\}$ is independent of the parameters.

There are two support points for each univariate D-optimal design (White, 1975) in $\mathbf{Z}$ and in $\mathbf{U}$. Let $\underline{z}$ and $\overline{z}$ denote these support points in $\mathbf{Z}$ and let $\underline{u}$ and $\overline{u}$ denote these supports in $\mathbf{U}$. Now, for $\mathbf{z} = (\underline{\mathbf{z}}, \overline{\mathbf{z}})^{\mathbf{T}}$ and $\mathbf{u} = (\underline{\mathbf{u}}, \overline{\mathbf{u}})^{\mathbf{T}}$, (5.2) can be written as

$$\det\{\mathbf{I}[\xi(\mathbf{z},\mathbf{u})]\} = \frac{1}{16}w(\underline{z})w(\overline{z})(\overline{z}-\underline{z})^2 w(\underline{u})w(\overline{u})(\overline{u}-\underline{u})^2. \tag{5.3}$$

Because the cumulative distribution function Φ is symmetric, it follows from invariance arguments that the D-optimal design will be symmetric about the origin. Thus $\underline{z} = -\overline{z}$,

$\underline{u} = -\overline{u}$, $\overline{z} > 0$, $\overline{u} > 0$, and $(\overline{z}, \overline{u})$ must maximize $w^2(z)z^2w^2(u)u^2$ over $(0, \infty) \times (0, \infty)$. As expected from White (1975), using the Splus function *nlmin,* the maximum is obtained at $\overline{z} = \overline{u} = 1.138$. It follows that D-optimal designs for two independent drugs with two associated independent responses include

- two standardized design points, (1.138,1.138), (-1.138,-1.138) with 50% of the trials at each point;
- two standardized design points, (-1.138,1.138), (1.138,-1.138) with 50% of the trials at each point;
- four standardized design points, (1.138,1.138), (-1.138,-1.138), (-1.138,1.138), and (1.138,-1.138) with 25% of the trials at each of them.

The determinant of the normalized information matrix, for any one of these designs is 0.0394752. That these maxima are globally D-optimal is verified using the general equivalence theorem. Figure 1 is the graph of the

$$trace\left\{\mathbf{I}^{-1}\left[\xi\left(\mathbf{z},\mathbf{u}\right);\theta\right] \times \mathbf{I}\left[\left(\mathbf{z_i},\mathbf{u_j}\right);\theta\right]\right\}$$

for the four point optimal design. Figure 2 graphs this trace on the circle that inscribes the design points. It can be seen in Figure 2 that trace$\{\mathbf{I}^{-1}\left[\xi\left(\mathbf{z},\mathbf{u}\right);\theta\right] \times \mathbf{I}\left[\left(\mathbf{z_i},\mathbf{u_j}\right);\theta\right]\} \leq 4$ for all (z_i, u_j) with equality when $(z_i, u_j) \in (\mathbf{z}, \mathbf{u})$.

In the case where $\sigma_1 = \sigma_2$ we have the following result.

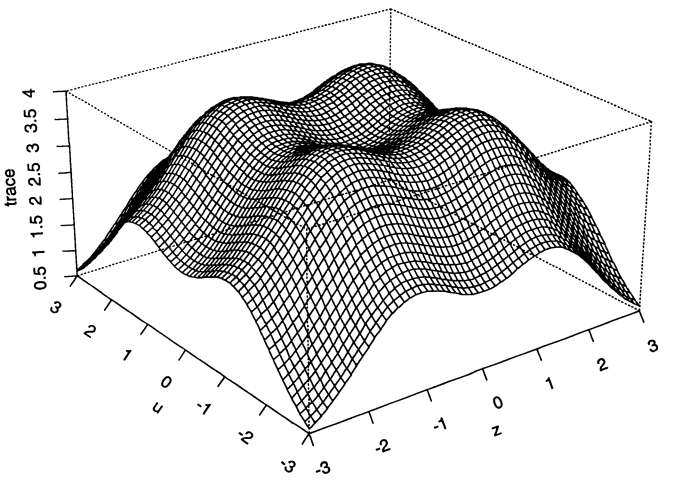

FIG. 1. *Trace in two-dimensional standardized space*

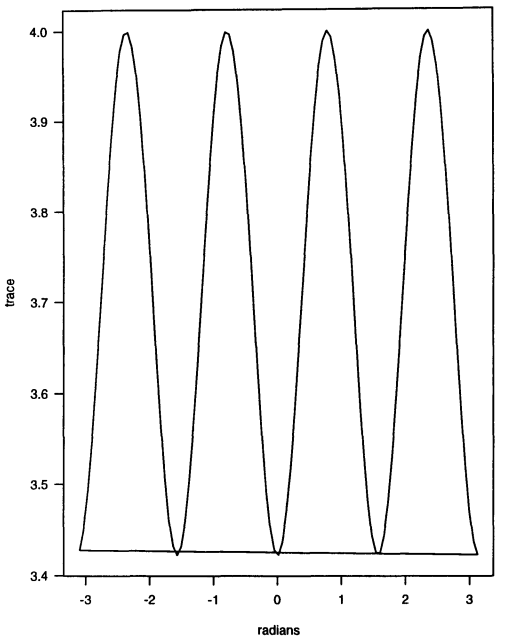

FIG. 2. *Trace over the circle that inscribes the D-optimal design points*

THEOREM 5.1 *When responses follow independent probit functions with* $\sigma_1 = \sigma_2 = \sigma$, *D-optimal designs include two standardized points* $(0.937, 0.937)$, $(-0.937, -0.937)$, *the two standardized points* $(0.937, -0.937)$, $(-0.937, 0.937)$ *and the four standardized points* $(0.937, 0.937)$, $(-0.937, -0.937)$, *(0.937,-0.937) and* $(-0.937, 0.937)$ *with equal allocation to each point in each design.*

PROOF: In this case, $\theta = (\mu_1, \mu_2, \sigma)^T$ and Fisher's information matrix is

$$\mathbf{I}[\xi(\mathbf{x}); \theta] = \frac{\mathbf{n}}{\sigma^2} \sum_{\mathbf{i}} \sum_{\mathbf{j}} \xi_{\mathbf{ij}} \begin{bmatrix} w(z_i) & 0 & z_i w(z_i) \\ 0 & w(z_j) & z_j w(z_j) \\ z_i w(z_i) & z_j w(z_j) & z_i^2 w(z_i) + z_j^2 w(z_j) \end{bmatrix}. \tag{5.4}$$

Consider four points $(\mathbf{z}, \mathbf{u}) = [(z, u), (-z, u), (z, -u), (-z, -u)]$ in $\mathbf{Z} \times \mathbf{U}$. Because these points are symmetric, (5.4) becomes

$$\mathbf{I}[\xi(\mathbf{z}, \mathbf{u})] = \frac{\mathbf{1}}{\mathbf{2}} \begin{bmatrix} w(z) & 0 & 0 \\ 0 & w(u) & 0 \\ 0 & 0 & z^2 w(u) + u^2 w(u) \end{bmatrix}.$$

It follows that

$$\det\{\mathbf{I}[\xi(\mathbf{z}, \mathbf{u})]\} = w(z) \times w(u) \times \left[\overline{z}^2 w(\overline{u}) + \overline{u}^2 w(\overline{u})\right].$$

A graph of the determinant of the Fisher information matrix over $(0, 3) \times (0, 3)$ shows it to be unimodal. Using the Splus function *nlmin,* the maximum is obtained at $z =$

$u = 0.937$. D-optimal designs in this case include the two point design, $(0.937, 0.937)$, $(-0.937, -0.937)$, the two point design $(0.937, -0.937)$, $(0.937, -0.937)$ and a four point design $(0.937, 0.937)$, $(-0.937, -0.937)$, (-0.937,0.937), (0.937,-0.937), with equal allocation to each point in each design. The determinant of the normalized information matrix for any one of these designs is equal to 0.170315.

To verify that the maximum is globally D-optimal, the general equivalence theorem can be used again. Figures like Figure 1 and Figure 2 were drawn,but are not shown, to show that trace$\{\mathbf{I}^{-1}[\xi(\mathbf{z},\mathbf{u});\theta] \times \mathbf{I}[(\mathbf{z_i},\mathbf{u_j});\theta]\} \leq 3$, for all (z_i, u_j) with equality when $(z_i, u_j) \in (\mathbf{z},\mathbf{u})$. □

6. Synergistic drug action. When responses follow a bivariate probit function, the response probabilities are

$$
\begin{aligned}
F_{11}[(x_i,x_j);\theta] &= \Phi\left(\frac{x_i-\mu_1}{\sigma_1}, \frac{x_j-\mu_2}{\sigma_2}\right) \\
F_{10}[(x_i,x_j);\theta] &= \Phi\left(\frac{x_i-\mu_1}{\sigma_1}\right) - \Phi\left(\frac{x_i-\mu_1}{\sigma_1}, \frac{x_j-\mu_2}{\sigma_2}\right) \\
F_{01}[(x_i,x_j);\theta] &= \Phi\left(\frac{x_j-\mu_2}{\sigma_2}\right) - \Phi\left(\frac{x_i-\mu_1}{\sigma_1}, \frac{x_j-\mu_2}{\sigma_2}\right)
\end{aligned} \tag{6.1}
$$

with

$$
F_{00}[(x_i,x_j);\theta] = 1 - \Phi\left(\frac{x_i-\mu_1}{\sigma_1}\right) - \Phi\left(\frac{x_j-\mu_2}{\sigma_2}\right) + \Phi\left(\frac{x_i-\mu_1}{\sigma_1}, \frac{x_j-\mu_2}{\sigma_2}\right)
$$

where $\Phi(z_i, z_j)$ is the standard bivariate normal cumulative distribution function, $\Phi(z)$ is the standard univariate cumulative normal distribution function and $\theta = (\mu_1, \sigma_1, \mu_2, \sigma_2, \rho)^{\mathbf{T}}$.

Fisher's information matrix is

$$
\mathbf{I}(\xi;\theta) = \mathbf{n}\sum_{\mathbf{i}=1}^{\mathbf{r}}\sum_{\mathbf{j}=1}^{\mathbf{c}}\xi_{\mathbf{ij}}\begin{pmatrix} \mathbf{I_{11}}(\mathbf{z_i},\mathbf{u_j}) & \mathbf{I_{12}}(\mathbf{z_i},\mathbf{u_j}) & \mathbf{I_{13}}(\mathbf{z_i},\mathbf{u_j}) \\ \mathbf{I_{12}^T}(\mathbf{z_i},\mathbf{u_j}) & \mathbf{I_{22}}(\mathbf{z_i},\mathbf{u_j}) & \mathbf{I_{23}}(\mathbf{z_i},\mathbf{u_j}) \\ \mathbf{I_{13}^T}(\mathbf{z_i},\mathbf{u_j}) & \mathbf{I_{23}^T}(\mathbf{z_i},\mathbf{u_j}) & \mathbf{I_{33}}(\mathbf{z_i},\mathbf{u_j}) \end{pmatrix},
$$

and

$$
\mathbf{I_{11}}(\mathbf{z_i},\mathbf{u_j}) = \begin{pmatrix} \frac{1}{\sigma_1^2}K(z_i,u_j) & \frac{z_i}{\sigma_1^2}K(z_i,u_j) \\ \frac{z_i}{\sigma_1^2}K(z_i,u_j) & \frac{z_i^2}{\sigma_1^2}K(z_i,u_j) \end{pmatrix}, \quad \mathbf{I_{23}}(\mathbf{z_i},\mathbf{u_j}) = \begin{pmatrix} \frac{1}{\sigma_2}U(z_i,u_j) \\ \frac{u_j}{\sigma_2}U(z_i,u_j) \end{pmatrix},
$$

$$
\mathbf{I_{22}}(\mathbf{z_i},\mathbf{u_j}) = \begin{pmatrix} \frac{1}{\sigma_2^2}S(z_i,u_j) & \frac{u_i}{\sigma_2^2}S(z_i,u_j) \\ \frac{u_i}{\sigma_2^2}S(z_i,u_j) & \frac{u_i^2}{\sigma_2^2}S(z_i,u_j) \end{pmatrix}, \quad \mathbf{I_{13}}(\mathbf{z_i},\mathbf{u_j}) = \begin{pmatrix} \frac{1}{\sigma_1}R(z_i,u_j) \\ \frac{z_i}{\sigma_2}R(z_i,u_j) \end{pmatrix},
$$

$$
\mathbf{I_{12}}(\mathbf{z_i},\mathbf{u_j}) = \begin{pmatrix} \frac{1}{\sigma_1\sigma_2}P(z_i,u_j) & \frac{u_i}{\sigma_1\sigma_2}P(z_i,u_j) \\ \frac{z_i}{\sigma_1\sigma_2}P(z_i,u_j) & \frac{z_iu_j}{\sigma_1\sigma_2}P(z_i,u_j) \end{pmatrix}, \quad \mathbf{I_{33}}(\mathbf{z_i},\mathbf{u_j}) = \mathbf{V}(\mathbf{z_i},\mathbf{u_j}),
$$

and the functions K, P, S, R, U and V, respectively, are defined by

$$K(z_i, u_j) = \left(\frac{\partial F_{11}}{\partial z_i}\right)^2 \left(\frac{1}{F_{11}} + \frac{1}{F_{01}}\right) + \left[\phi(z_i) - \frac{\partial F_{11}}{\partial z_i}\right]^2 \left(\frac{1}{F_{10}} + \frac{1}{F_{00}}\right),$$

$$\begin{aligned} P(z_i, u_j) \;=\; & \frac{1}{F_{11}} \left(\frac{\partial F_{11}}{\partial z_i}\right) \left(\frac{\partial F_{11}}{\partial u_j}\right) - \frac{1}{F_{10}} \left(\frac{\partial F_{11}}{\partial u_j}\right) \left[\phi(z_i) - \frac{\partial F_{11}}{\partial z_i}\right] \\ & - \frac{1}{F_{01}} \left(\frac{\partial F_{11}}{\partial z_i}\right) \left[\phi(u_j) - \frac{\partial F_{11}}{\partial u_j}\right] + \frac{1}{F_{00}} \left[\phi(z_i) - \frac{\partial F_{11}}{\partial z_i}\right] \\ & \times \left[\phi(u_j) - \frac{\partial F_{11}}{\partial u_j}\right], \end{aligned}$$

$$S(z_i, u_j) = \left(\frac{\partial F_{11}}{\partial u_j}\right)^2 \left(\frac{1}{F_{11}} + \frac{1}{F_{10}}\right) + \left[\phi(u_j) - \frac{\partial F_{11}}{\partial u_j}\right]^2 \left(\frac{1}{F_{01}} + \frac{1}{F_{00}}\right),$$

$$R(z_i, u_j) = -\left(\frac{\partial F_{11}}{\partial z_i}\right) \left(\frac{1}{F_{11}} + \frac{1}{F_{10}} + \frac{1}{F_{01}} + \frac{1}{F_{00}}\right) + \phi(z_i) \left(\frac{1}{F_{10}} + \frac{1}{F_{00}}\right),$$

$$U(z_i, u_j) = -\left(\frac{\partial F_{11}}{\partial u_j}\right) \left(\frac{1}{F_{11}} + \frac{1}{F_{10}} + \frac{1}{F_{01}} + \frac{1}{F_{00}}\right) + \phi(u_j) \left(\frac{1}{F_{01}} + \frac{1}{F_{00}}\right),$$

and

$$V(z_i, u_j) = \left(\frac{1}{F_{11}} + \frac{1}{F_{10}} + \frac{1}{F_{01}} + \frac{1}{F_{00}}\right).$$

With synergy of drug action in the bivariate probit response function, there is no transformation that will eliminate the dependency of the design on ρ. Because of computational difficulties we only considered designs of the form $\xi(\mathbf{z}, \mathbf{u})$, where

$$(\mathbf{z}, \mathbf{u}) = [(z, z), (-z, z), (z, -z), (-z, -z)],$$

with $z \in R$. We obtain the design points for which the determinant of the 5 parameter Fisher information matrix is maximum as a function of the correlation coefficient. For $\rho = 0, 0.1, 0.5, 0.90$, we plotted $\det\{\mathbf{I}[\xi(\mathbf{z}, \mathbf{u}); \theta]\}$, $-3 \leq z, u \leq 3$. These functions were all unimodal and their modes were identified using Splus functions *nlminb*. The points at which the maxima is obtained are given in Table 1. In practice there are many ways to use optimal designs depending on the goal of the experiment. One way to build in robustness is to sequentially construct the D-optimal designs. For example one might start the design with a best guess for $(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$, then sequentially estimate the parameters and adapt the design points accordingly. We hope that our preliminary

TABLE 1

D-optimal design in standardized space for several degrees of synergism

$\rho = 0$	$\rho = 0.1$	$\rho = 0.5$	$\rho = 0.90$
$(1.138, 1.138)$	$(0.94, 0.94)$	$(0.89, 0.89)$	$(0.939, 0.939)$
$(-1.138, 1.138)$	$(-0.94, 0.94)$	$(-0.89, 0.89)$	$(-0.939, 0.939)$
$(1.138, -1.138)$	$(0.94, -0.94)$	$(0.89, -0.89)$	$(0.939, -0.939)$
$(-1.138, -1.138)$	$(-0.94, -0.94)$	$(-0.89, -0.89)$	$(-0.939, -0.939)$

results for bivariate probit response functions will stimulate further development of this useful model.

Acknowledgements. The author would like to thank the referees for their most valuable comments and suggestions which have significantly improved this paper. The author also would like to thank Professor Nancy Flournoy, his dissertation advisor, for her advice and support.

REFERENCES

ASHFORD, J. R. (1981). General models for joint action of mixtures of drugs. *Biometrics* **37** 457-474.

ASHFORD, J. R. AND COBBY, J. M. (1974). A system of models for action of drugs applied singly or jointly to biological organisms. *Biometrika* **30** 11-31.

ASHFORD, J. R. AND SMITH, C. S. (1964). General models for quantal response to the joint action of a mixture of drugs. *Biometrika* **51** 413-428.

ASHFORD, J. R. AND SMITH, C. S. (1965). An alternative system for the classification of mathematical models for quantal responses to mixtures of drugs in biological assay. *Biometrics* **21** 181-188.

ASHFORD, J. R. AND SOWDEN, R. R. (1970). Multivariate probit analysis. *Biometrics* **26** 535-615.

ATKINSON, A. C. (1996). The usefulness of optimum experimental designs. *J. Roy. Statist. Soc. B* **58** 59-76.

ATKINSON, A.C. AND DONEV, A. N. (1992). Optimum Experimental Designs. Clarendon Press, Oxford.

CHALONER, K. AND VERDINELLI, I. (1995). Bayesian experimental design: A review. *Statist. Sci.* **10** 273-304.

CHERNOFF, H. (1953). Locally optimal designs for estimating parameters. *Ann. Math. Statist.* **24** 586-602.

COCHRAN, W. G. (1973). Experiments for nonlinear functions. *J. Am. Statist. Assoc.* **68** 771-781.

DERMAN, C. (1957). Nonparametric up and down experimentation. *Ann. Math. Statist.* **28** 795-798.

DURHAM, S. D. AND FLOURNOY, N. (1994). Random walks for quantile estimation. In *Statistical Decision Theory and Related Topics V* (S. S. Gupta and J. O. Berger, eds.), 467-476. Springer-Verlag, New York.

DURHAM, S. D. AND FLOURNOY, N. (1995). Up-and-down designs I: Stationary treatment distributions. In *Adaptive Designs* (N. Flournoy and W. F. Rosenberger, eds.), 139-157. Institute of Mathematical Statistics, Hayward.

FEDOROV, V. V. (1972). *Theory of Optimal Experiments.* Academic Press, New York.

FEDOROV, V. V. (1981). Active regression experiments. In *Mathematical Methods of Experimental Design,* 19-73, Nauka Press, Siberian Section (in Russian).

FLOURNOY, N. (1993). A clinical experiment in bone marrow transplantation: estimating a percentage point of a quantal response curve. In *Case Studies in Bayesian Statistics* (C. Gatsonis, J. Hodges, R. E. Kass and N. Singpurwalla, eds.), 324-336. Springer-Verlag, New York.

FORD, I., TORSNEY, B. AND WU, C. F. J. (1992). The use of canonical form in the construction of locally optimal designs for non-linear problems. *J. Roy. Statist. Soc. B* **54** 569-583.

FORD, I., TITTERINGTON, D. M. AND KITSOS, C. P.(1989). Recent advances in non-linear experimental designs. *Technometrics* **31** 49-60.

HALL, J. L. (1968). Maximum likelihood sequential procedure for estimation of psychometric functions. *J. Acoust. Soc. Amer.* **44** 370.

GIOVAGNOLI, A. AND PINTACUDA, N. (1996). Discussion of A. C. Atkinson: The usefulness of optimum experimental designs. *J. Roy. Statist. Soc. B* **58** 98.

GIOVAGNOLI, A. AND PINTACUDA, N. (1998). Properties of frequency distributions induced by general "up-and-down" methods for estimating quantiles. *J. Statist. Plann. Inf.*, in press.

HEISE, M. A. AND MYERS, R. M. (1996). Optimal designs for bivariate logistic regression. *Biometrics* **52** 613-624.

HEWLETT, P. S. AND PLACKETT, R. L. (1959). A unified theory for quantal responses to mixtures of drugs: non-interactive action. *Biometrics* **15** 591 – 610.

HEWLETT, P. S. AND PLACKETT, R. L. (1964). A unified theory for quantal responses to mixtures of drugs: competitive action. *Biometrics* **20** 556 – 575.

HEWLETT, P. S. AND PLACKETT, R. L. (1979). *The Interpretation of Quantal Responses in Biology.* Edward Arnold, London.

KIEFER, J. AND WOLFOWITZ, J. (1960). The equivalence of two extremum problems. *Can. J. Math.* **12** 363-366.

PLACKETT, R. L. AND HEWLETT, P. S. (1963). A unified theory for quantal reponses to mixtures of drugs; the fitting to data of certain models for two non-interactive drugs with complete positive correlation of tolerance. *Biometrics* **19** 517-531.

PLACKETT, R. L. AND HEWLETT, P. S. (1952). Quantal responses to mixtures of poison. *J. Roy. Statist. Soc. B* **14** 141 – 163.

PLACKETT, R. L. AND HEWLETT, P. S. (1967). A comparison of two approaches to the construction of models for quantal responses to mixtures of drugs. *Biometrics* **23** 27-44.

SILVEY, S. D. (1980). *Optimal Design.* Chapman & Hall, London.

SITTER, R. R. AND TORSNEY, B. (1995). Optimal designs for binary response experiments with two design variables. *Statist. Sinica* **5** 405-419.

TSUTAKAWA, R. K. (1980). Design of dose levels for estimating a percentage point of a logistic quantal response curve. *Appl. Statist.* **29** 25-33.

WHITE, L. (1975). *Optimal Design of Experiments for Estimation in Nonlinear Models.* Imperial College of Science and Technology, London (doctoral thesis).

CLINICAL TRIALS & SURVEYS CORP.
350 WEST QUADRANGLE
BALTIMORE, MD 21210 U.S.A.