

## *Chapter 6*

### **MODEL ROBUST BAYESIAN DESIGNS**

#### **§6.1 Motivation**

Three popular models used in toxicity studies include the Poisson exponential model, the Poisson linear model, and the square root transformation model discussed in Chapter 1. The optimal experimental design for the Poisson exponential model is quite different from the one for the Poisson linear and square root transformation models. Since the optimal design is model dependent, the researcher must have some knowledge about the true form of the model in order to select the best design. This model selection process is relatively easy for the experimenter who has past research that consistently indicates an appropriate model. Yet, in many cases the answer is not so clear. The researcher is faced with a dilemma if previous studies suggest that two or more models have fit well in similar experimental situations. An even greater challenge arises in the search for the optimal design when researchers begin experimentation with new toxicants or recently developed drugs where there are no past studies. Regardless of the circumstances, the use of an experimental design based on the wrong model could result in a prodigious waste of time and resources. Thus, the need arises for experimental designs which are robust in the presence of model uncertainty. In

linear models, this need has been addressed briefly by Dumouchel and Jones (1994) and Neff (1997). This chapter focuses on the development of useable model robust Bayesian D-optimal designs for the aforementioned models.

## §6.2 The Models

The three models used in the construction of the model robust Bayesian D-optimal designs are listed in Table 6.2.1. The table includes the functional form, the expected value at a design point  $i$ , and the expected value at the control.

Table 6.2.1 Models for Bayesian D-Optimal Model Robust Designs

<i>Name</i>	<i>Form</i>	<i>E(point i)</i>	<i>E(control)</i>
Poisson Exponential	$y_{ij} = e^{\beta_0 + \beta_1 x_i} + \epsilon_{ij}$	$e^{\beta_0 + \beta_1 x_i}$	$e^{\beta_0}$
Poisson Linear	$y_{ij} = \beta_0 + \beta_1 x_i + \epsilon_{ij}$	$\beta_0 + \beta_1 x_i$	$\beta_0$
Square Root Transformation	$\sqrt{y_{ij}} = \beta_0 + \beta_1 x_i + \epsilon_{ij}$	$\beta_0 + \beta_1 x_i$	$\beta_0$

## §6.3 D-Optimality Criterion for the Non-Bayesian Design

Recall that the D-optimal design requires the maximization of the determinant of the information matrix of the coefficients. This is equivalent to minimizing their generalized variance. The general theory for D-optimality as well as the derivation of the two level D-optimal design for the Poisson exponential model appears in Chapter 2. The D-optimality criterion for the other two models used in the development of the model robust design are addressed in the following sections.

The first designs discussed are the two level designs. Two level designs are often considered impractical by some researchers because they provide no way of testing for lack of fit. Other practitioners are of the opinion that two level designs are useful in the single regressor case when one is very confident in his or her knowledge of  $\lambda_c$ . Recall that  $\lambda_c$  is the expected value of the response in the presence of no toxicant so  $\lambda_c = f(0, \beta) = g(\beta_0)$ . In other words, the response at the control is solely a function of  $b_0$  as shown in Table 6.2.1. Since it is very easy to gain information about the response in the absence of toxicant, this often eliminates the need for estimation of  $b_0$ .

This implies that a two level design would be sufficient for the estimation of  $\beta_1$  as well as lack of fit if  $\beta_0$  is not estimated.

Regardless of their practicality, one cannot argue against the theoretical importance of these designs. Using equivalence theory, Chiacchierini (1996) showed that two-level designs are indeed optimal out of the class of all D-optimal designs. Hence, they serve as a benchmark by which to judge all other D-optimal designs. For this reason two level designs play an important role in Bayesian D-optimal designs.

### §6.3.1 The Two Level Design for the Square Root Transformation Model

The determinant of the information matrix for the square root transformation model is

$$|\mathbf{I}_T| = \left( \frac{1}{\beta_1^2} \right) (n_1 n_2 (\sqrt{\lambda_1} - \sqrt{\lambda_2})^2). \quad (6.3.1)$$

Of course, the criterion is a function of its unknown parameters. From this form, it can be transformed into a function of effective concentrations by the substitution  $\lambda_i = q_i \lambda_c$ . So the determinant becomes

$$|\mathbf{I}_T| = \left( \frac{\lambda_c}{\beta_1^2} \right) (n_1 n_2 (\sqrt{q_1} - \sqrt{q_2})^2). \quad (6.3.2)$$

The right hand side of (6.3.2) is obviously maximized when  $n_1 = n_2 = n$ . The terms  $\lambda_c$  and  $\frac{1}{\beta_1^2}$  are treated as constants with respect to the maximization. Now, the values of  $q_i$  must be chosen so that factor  $(\sqrt{q_1} - \sqrt{q_2})^2$  is maximized. This quantity will be maximized when the distance between  $q_1$  and  $q_2$  is as large as possible. Thus, the D-optimal design for the square root transformation model is accomplished by placing 50% of the observations at the  $EC_0$  and the remaining 50% are placed at the  $EC_{100}$ . The general concepts seen in the development of this are present in the development of the optimal designs for the Poisson linear model in the next section.

### §6.3.2 The Two Level Design for the Poisson Linear Model

The D-optimality criterion for the two level design for the Poisson linear model is

$$|I_L| = \left( \frac{1}{\beta_1^2} \right) \left( \frac{n_1 n_2 (\lambda_1 - \lambda_2)^2}{\lambda_1 \lambda_2} \right). \quad (6.3.3)$$

Using the substitution of  $\lambda_i = q_i \lambda_c$  to make the criterion independent of its parameters yields

$$|I_L| = \left( \frac{1}{\beta_1^2} \right) \left( \frac{n_1 n_2 (q_1 - q_2)^2}{q_1 q_2} \right). \quad (6.3.4)$$

In order to maximize this function, the numerator should be made as large as possible and the denominator should be made as small as possible. Since the numerator is a function of the distance between  $q_1$  and  $q_2$ , it will be maximized when 50% of the observations are placed at the  $EC_0$  and the remaining 50% are placed at the  $EC_{100}$ . The denominator, in turn, should be made as small as possible. This condition is also met when the design points are the  $EC_0$  and the  $EC_{100}$ . However, this choice of design points would make the denominator zero and the value of the determinant undefined. Hence the optimal design will have 50% of the observations at the  $EC_{100}$  and 50% at a point as close to the  $EC_0$  as possible (Chiacchierini, 1996).

### §6.4 D-Optimality Criterion for the Three Level Designs

As previously mentioned, two level designs are of questionable practical value to the biologist. This is especially true in the presence of model uncertainty. For this reason, only three level model robust Bayesian designs will be presented. Since the three level D-optimality criteria for the individual models plays an important role in the Bayesian designs, the criteria are detailed below. Chiacchierini derived expressions for the determinant of the information matrix for both Poisson models. The expression for the determinant of the information matrix for the Poisson exponential model is

$$\begin{aligned}
|\mathbf{I}_E| &= \left( \frac{1}{\boldsymbol{\beta}_1^2} \right) \left( n_1 n_2 \lambda_1 \lambda_2 \left( \ln \frac{\lambda_1}{\lambda_2} \right)^2 + n_1 n_3 \lambda_1 \lambda_3 \left( \ln \frac{\lambda_1}{\lambda_3} \right)^2 + n_2 n_3 \lambda_2 \lambda_3 \left( \ln \frac{\lambda_2}{\lambda_3} \right)^2 \right) \\
&= \left( \frac{\lambda_c}{\boldsymbol{\beta}_1^2} \right) \left( n_1 n_2 q_1 q_2 \left( \ln \frac{q_1}{q_2} \right)^2 + n_1 n_3 q_1 q_3 \left( \ln \frac{q_1}{q_3} \right)^2 + n_2 n_3 q_2 q_3 \left( \ln \frac{q_2}{q_3} \right)^2 \right).
\end{aligned} \tag{6.4.1}$$

The D-optimality criterion for the Poisson Linear model is given by

$$\begin{aligned}
|\mathbf{I}_L| &= \left( \frac{1}{\boldsymbol{\beta}_1^2} \right) \left( n_1 n_2 \frac{(\lambda_1 - \lambda_2)^2}{\lambda_1 \lambda_2} + n_1 n_3 \frac{(\lambda_1 - \lambda_3)^2}{\lambda_1 \lambda_3} + n_2 n_3 \frac{(\lambda_2 - \lambda_3)^2}{\lambda_2 \lambda_3} \right) \\
&= \left( \frac{1}{\boldsymbol{\beta}_1^2} \right) \left( n_1 n_2 \frac{(q_1 - q_2)^2}{q_1 q_2} + n_1 n_3 \frac{(q_1 - q_3)^2}{q_1 q_3} + n_2 n_3 \frac{(q_2 - q_3)^2}{q_2 q_3} \right).
\end{aligned} \tag{6.4.2}$$

Finally, the expression to be maximized in the square root transformation model is

$$\begin{aligned}
|\mathbf{I}_T| &= \left( \frac{1}{\boldsymbol{\beta}_1^2} \right) \left( n_1 n_2 (\sqrt{\lambda_1} - \sqrt{\lambda_2})^2 + n_1 n_3 (\sqrt{\lambda_1} - \sqrt{\lambda_3})^2 + n_2 n_3 (\sqrt{\lambda_2} - \sqrt{\lambda_3})^2 \right) \\
&= \left( \frac{\lambda_c}{\boldsymbol{\beta}_1^2} \right) \left( n_1 n_2 (\sqrt{q_1} - \sqrt{q_2})^2 + n_1 n_3 (\sqrt{q_1} - \sqrt{q_3})^2 + n_2 n_3 (\sqrt{q_2} - \sqrt{q_3})^2 \right).
\end{aligned} \tag{6.4.3}$$

In all cases, notice the similarity of the three level criterion to its two level counterpart. With the three level criteria specified, these expressions can now be combined to form a model robust Bayesian D-optimal design criterion.

### §6.5 Model Robust Optimal Bayesian Designs

The topic of model robust design lends itself naturally to a Bayesian perspective. By placing a discrete prior on the models under consideration and minimizing the Bayes risk, a design which performs reasonably well for all of these models can be obtained. The general formula for the Bayes design in this situation is given by the following expression

$$\text{Max}_{\delta \in \mathcal{D}} \sum_i R(\delta, M_i) P(M_i) \quad (6.5.1)$$

where  $M_i$  represents the  $i^{\text{th}}$  model,  $P(M_i)$  is the prior density on the models under consideration,  $R(\delta, M_i)$  is the design optimality criterion of choice, and  $d$  is any design from the set of all possible designs  $\mathcal{D}$ . The quantity  $R(\delta, M_i)$  is the negative of the Bayes risk so by maximizing this function, the Bayes risk is minimized. Note that this expression is the discrete form of the Bayesian design criterion 2.5.1 in Chapter 2.

Since  $R(\delta, M_i)$  is a design criterion, it seems that one would proceed by simply substituting the determinants for their respective  $R(\delta, M_i)$ . However, several complications prohibit this from being the correct approach. The first complication arises from the forms of the determinants in Section 6.4 when the substitution  $\lambda_i = q_i \lambda_c$  is made. A glance at (6.4.1), (6.4.2), and (6.4.3) reveal that each of the three criteria is a function of a different power of  $\lambda_c$ . The Poisson exponential model is function of  $\lambda_c^2$ , the linear model is independent of  $\lambda_c$ , and the square root transformation model is a function of  $\lambda_c$ . Since this value cannot be treated as a constant with respect to maximization, a different design must be generated for each value of  $\lambda_c$ . This quality of the criterion is undesirable since it does not allow for the formulation of a general design for each prior.

A second problem resulting from this approach lies in the scale of the D-optimal criterion for each of the different models considered. The determinant of the information matrix has its origins in the log likelihood. The log likelihoods for different models are on different scales. Thus, the Bayesian D-optimality criterion is a linear combination of determinants on different scales. The resulting design would not be robust to all three models and would only be optimal for the particular model that dominated the criterion. So an equivalent Bayesian D-optimality criterion must be developed that is invariant to both scale and the choice of  $\lambda_c$ .

A scale invariant Bayesian D-optimality criterion is accomplished by measuring the performance of each design as a proportion with respect to an optimal design for that model. Note

that this also eliminates the dependency of the individual model criteria on different powers of  $\lambda_c$ . The new formula is given below

$$R(\delta, M_i) = \frac{|\mathbf{I}(\delta, M_i)|}{\max_{\delta \in \mathcal{D}} |\mathbf{I}(\delta, M_i)|}. \quad (6.5.2)$$

The value of the denominator, the optimal determinant, will change depending on the restrictions placed on the design. Hence, the Bayesian design optimality criterion becomes a weighted average of “efficiencies” where the weights are the prior probabilities associated with the models under consideration. The expansion of (6.5.1) using (6.5.2) as  $R(d, M_i)$  gives the following form for the Bayesian criteria:

$$\frac{|\mathbf{I}(\delta, M_E)|}{\max_{\delta \in \mathcal{D}} |\mathbf{I}(\delta, M_E)|} P(M_E) + \frac{|\mathbf{I}(\delta, M_L)|}{\max_{\delta \in \mathcal{D}} |\mathbf{I}(\delta, M_L)|} P(M_L) + \frac{|\mathbf{I}(\delta, M_T)|}{\max_{\delta \in \mathcal{D}} |\mathbf{I}(\delta, M_T)|} P(M_T). \quad (6.5.2)$$

With the criterion defined, the next step in finding model robust Bayesian designs is to define the priors for the models of interest.

### §6.6 Priors on the Models

While any number of discrete priors could be specified for the three models in question, five priors were chosen for the subsequent designs. These priors range from completely uncertain to very confident in their assignment of probabilities to the models. The Nelder-Mead algorithm is then used to find the optimal three level design based on the resulting criterion.

Table 6.6.1 List of Priors for the Models.

Prior	$P(M_E)$	$P(M_L)$	$P(M_T)$
1	0.3	0.3	0.3
2	0.6	0.3	0.1
3	0.8	0.15	0.05
4	0.15	0.8	0.05
5	0.05	0.15	0.8

### §6.7 Three Level Bayesian D-Optimal Model Robust Designs

The first effort to generate three level designs was done through the development of unrestricted designs. Unrestricted designs, as the name implies, have no restrictions placed on the allocation of experimental units to design points or the design points themselves. The appropriate substitutions were made for the priors and  $R(\delta, M_i)$  for three level criteria as shown in expressions (6.5.1) and in its expanded form in expression (6.5.3). The denominator used in the function  $R(\delta, M_i)$  is the value obtained by substituting the optimal design levels and allocation percentages for the two level designs detailed in Section 2.4 of Chapter 2 and Sections 6.3.1 and 6.3.2.

Designs were generated for several regions of operability including the  $[EC_{0.05}, EC_{100}]$ ,  $[EC_{20}, EC_{80}]$ ,  $[EC_{30}, EC_{70}]$ , and  $[EC_{10}, EC_{50}]$ . The Nelder-Mead algorithm was used for optimization. For the five priors and the four regions of operability, the three level design criterion selected two level optimal Bayesian designs. Since the resulting two-level designs are not practical in the presence of model uncertainty, they will not be detailed here. Thus, the need for restricted designs is motivated by the undesirable qualities of unrestricted designs as was done in Chapters 3 and 4.

The first of these restrictions imposed in an effort to find useable three level designs involves the control. Since biologists want to take observations at the control for any experiment, the third level is set at the control. The other two levels are free to vary provided that they are symmetric around the region of operability. The final constraint restricts the number of observations at the non-control design points to be equal. This last restriction places these designs in a class that is very popular with biologists. For these restricted designs, the value of the denominator of  $R(\delta, M_i)$  is based on the optimal restricted design found by Chiacchierini (1996). This is justified because one would want to select this design based on its performance with respect to a three level counterpart.

The model robust Bayesian designs for the region of interest of  $[EC_{20}, EC_{80}]$  appear in Table 6.7.1. Two types of efficiencies appear in the table. The two types of efficiencies in the table are the unrestricted and restricted efficiencies. The unrestricted efficiency involves a comparison between the Bayesian design and the two level optimal design on the corresponding region of



operability while the restricted efficiencies compare it to its restricted counterpart. Three efficiencies exist for each type because one must consider the efficiency of the design of interest against each of the three underlying models. This allows for consideration of the “worst case scenario” in the event that the true model is not the one receiving the highest prior probability.

Designs on the  $[EC_{20}, EC_{80}]$  are particularly impressive with regard to both regular and restricted efficiencies. In fact, this optimal restricted design differs from the optimal non-Bayesian designs for the three models in allocation of design points only. The levels are the same. Note the similarities in the percentages of allocation of design points coupled with the unity of design levels across priors. This seems to indicate that the design is robust to the prior as well. Three level designs are also found for other regions of operability including the  $[EC_{0.05}, EC_{100}]$ , the  $[EC_{30}, EC_{70}]$ , and the  $[EC_{10}, EC_{50}]$ . They are tabled in Appendix D. The comments made about the designs on the  $[EC_{20}, EC_{80}]$  are also valid about the design on the  $[EC_{30}, EC_{70}]$  and the  $[EC_{10}, EC_{50}]$ . In contrast, the design on the  $[EC_{0.05}, EC_{100}]$  does not possess the same high overall efficiencies and robustness to priors. In fact this design is only included for completeness since biologists are rarely interested in designs which cover such a large region.

### §6.8 Posterior Probabilities

As this chapter draws to a close, a comment on the use of posterior probabilities is necessary. A large part of Bayesian theory involves updating prior information with observed data to form posterior probabilities. The prior probabilities used in the design phase of the experiment can be combined with the data based likelihoods of the different models to form posterior probabilities. These posterior probabilities then function as a model selection criterion for the analysis of data collected from these designs. The formula for the posterior probabilities is

$$P(M_i | \mathbf{y}) = \frac{L(\mathbf{y} | M_i) P(M_i)}{\sum L(\mathbf{y} | M_i) P(M_i)}. \quad (6.8.1)$$

One should remember that the joint likelihoods must arise from the same distribution for this probability to be valid. This ensures that likelihoods are on the same scale. This technique is best

suited for GLM models when one is faced with making a choice among models based on different link functions. Preliminary case studies using the Poisson exponential and the Poisson linear models indicate that this method works well in identifying the true model.

## **§6.9 Conclusions**

Model robust D-optimal Bayesian designs perform very well in the presence of model uncertainty. In the majority of regions of interest examined efficiencies based on the optimal restricted design for all models are nearly 100%. Their robustness to the prior also makes them attractive. The natural progression to the use of posterior probabilities as a GLM model selection criterion is another salient feature of these designs. As a final note, both of these techniques seem to offer much promise when applied to other alphabetic design criteria and other models.

**Table 6.2.2. D-optimal Three Level Bayesian Designs on  $(EC_{20}, EC_{80})$**

<b><i>Prior</i></b>	<b><math>p_1=p_2</math></b>	<b><math>x_1</math></b>	<b><math>x_2</math></b>	<b><math>p_c</math></b>	<b><math>RE_{EXP}</math></b>	<b><math>E_{EXP}</math></b>	<b><math>RE_{LIN}</math></b>	<b><math>E_{LIN}</math></b>	<b><math>RE_{TRA}</math></b>	<b><math>E_{TRA}</math></b>
$(\overline{0.3}, \overline{0.3}, \overline{0.3})$	0.3630	$EC_{20}$	$EC_{80}$	0.2740	99.87%	86.15%	99.87%	75.17%	100%	87.07%
(0.6,0.3,0.1)	0.3568	$EC_{20}$	$EC_{80}$	0.2684	99.94%	86.21%	99.78%	75.10%	99.97%	87.05%
(0.8,0.15,0.05)	0.3507	$EC_{20}$	$EC_{80}$	0.2986	99.99%	86.25%	99.66%	75.01%	99.92%	87.00%
(0.15,0.8,0.05)	0.3750	$EC_{20}$	$EC_{80}$	0.2500	99.63%	85.94%	99.98%	72.25%	99.97%	87.04%
(0.05,0.15,0.8)	0.3775	$EC_{20}$	$EC_{80}$	0.2450	99.80%	86.09%	100%	86.09%	100%	87.07%