

Chapter 3

BAYESIAN D-OPTIMAL DESIGNS

§3.1 Motivation

Optimal designs for impaired reproduction studies are typically listed in the form of effective concentrations and the number of experimental units to allocate to those ECs. When it comes to translating the design into natural units, researchers find that the ECs are functions of the unknown parameters. While practitioners can use an educated guess for the parameter, there are better approaches. One of these is a Bayesian technique that allows for uncertainty in the guesses of these parameters by specifying a prior on them.

Some general comments are in order on the specification of priors for β_0 and β_1 . Since the parameter β_0 is tied to the maximum amount of reproduction in the absence of toxicant, accurate information about this parameter is often readily available. Yet, the value of the “slope” parameter, β_1 , is not so easily specified. In many cases, the biologist may not have sufficient insight about the toxicant to place an educated guess on its value. However, the biologist may be prepared to give a

range on a particular effective concentration (EC). The ability to determine this range suggests the idea of a Bayesian design criterion that incorporates this a priori belief. Fortunately, for the exponential model, this range on an EC can be translated into a prior on the slope. This chapter examines the use of the D-optimality criterion in conjunction with uniform and normal priors on the parameters for the single regressor exponential model to obtain Bayesian optimal designs.

§3.2 Uniform Priors on the Parameters

For the priors in this work, the parameters β_0 and β_1 will be considered independent and thus the joint prior will be formulated as a product of the individual priors on β_0 and β_1 . First, the focus is on the prior on β_0 . Recall from Chapter 1 that the exponential model is given by $y_{ij} = e^{\beta_0 + \beta_1 x_i} + \epsilon_{ij}$ and the expected value in the absence of toxicant is $\lambda_c = e^{\beta_0}$. One must develop a prior on β_0 through λ_c . Lower and upper limits can be placed on the expected response at control say, $\lambda_{c,L}$ and $\lambda_{c,U}$. Since it is very easy to obtain information about reproduction in the absence of toxicant, the biologist would have no trouble specifying this range. Thus, $\beta_0 \sim U(a, b)$ where $a = \ln \lambda_{c,L}$ and $b = \ln \lambda_{c,U}$.

Now, the prior on the slope will be addressed. This prior begins as a range on an EC. For example, the researcher believes that the EC_{100q} falls between the lower and upper boundaries of EC_L and EC_U respectively as expressed in natural units. As shown previously, $\lambda_i = q_i \lambda_c = e^{\beta_0 + \beta_1 x_i}$. Since $\lambda_c = e^{\beta_0}$, $q_i = e^{\beta_1 x_i}$. By substituting the appropriate value for q_i and $x_i = EC_L$ or $x_i = EC_U$, the equation, $q_i = e^{\beta_1 x_i}$, can be solved to obtain the respective bounds for a uniform prior on β_1 . So, $\beta_1 \sim U(c, d)$ where $c = \frac{\ln(q_i)}{EC_L}$ and $d = \frac{\ln(q_i)}{EC_U}$. Note that the uniform prior on β_1 can be derived irrespective of the value of q_i . However, the most information is typically known about the EC_{50} . The prior on the intercept and the slope are combined to form the joint prior shown in (3.2.1).

$$\pi(\beta) = \pi(\beta_0)\pi(\beta_1) = \frac{1}{(b-a)(d-c)} \quad (3.2.1)$$

§3.3 Two Level Bayesian D-Optimality Criterion under the Uniform Prior

Recall that D-optimality maximizes the determinant of the Fisher information matrix. For the two level design in the exponential model, Chiacchierini described the determinant as

$$|I| = n_1 n_2 e^{\beta_0 + \beta_1 x_1} e^{\beta_0 + \beta_1 x_2} (x_1 - x_2)^2. \quad (3.3.1)$$

The Bayesian design criterion for the prior on the parameters is given by

$$\max_{\delta \in \mathcal{D}} \int_{\beta} R(\delta, \beta) \pi(\beta) d\beta \quad (3.3.2)$$

where $R(\delta, \beta)$ is the D-optimality criterion and $\pi(\beta)$ is the uniform prior placed on β . At this time, the joint prior on both β_0 and β_1 will be used. It will be shown that the criterion is invariant to the prior on β_0 . Theoretically, this makes sense because the formula for the value of q_i in an EC does not involve β_0 . It also has a nice practical parallel since the researcher can easily obtain information about β_0 through what is known about the control.

Let $\beta_0 \sim U(a, b)$ and $\beta_1 \sim U(c, d)$. Hence, the Bayesian criterion, (3.3.2), becomes

$$\int_{\beta} R(\delta, \beta) \pi(\beta) d\beta = \frac{n_1 n_2 (x_1 - x_2)^2 (e^{2b} - e^{2a}) (e^{d(x_1+x_2)} - e^{c(x_1+x_2)})}{(x_1 + x_2)(b - a)(d - c)} \quad (3.3.3)$$

This demonstrates that maximization of the criterion is not dependent on the ranges set for β_0 since the portion of the criterion pertaining to β_0 , $\frac{(e^{2b} - e^{2a})}{(b - a)}$, becomes a constant with respect to the design points. Thus, the Bayesian criterion only involves taking the expected value over the parameter β_1 . By finding the values of x_1 , x_2 , n_1 , and n_2 which maximize (3.3.4).

$$\int_{\beta_1} R(\delta, \beta_1) \pi(\beta_1) d\beta_1 = \frac{n_1 n_2 (x_1 - x_2)^2 (e^{d(x_1+x_2)} - e^{c(x_1+x_2)})}{(d - c)(x_1 + x_2)}, \quad (3.3.4)$$

the optimal two level Bayesian design can be found. Note that the design criterion is no longer a function of unknown parameters. Rather, it is a function of ranges based on practical knowledge. In this way, it accounts for the variability around the guess placed on an EC. Design points for the Bayesian optimal designs can be reported in natural units or in terms of the EC_{100q} with the Bayesian criterion by using numerical optimization methods. The design points found in natural units can easily be converted to EC_{100q} by the following formula $q_i = e^{\frac{(c+d)}{2}x_i}$ and vice versa. Note that $\frac{(c+d)}{2}$ is the mean of the uniform distribution on β_1 .

§3.4 Ratio Constancy of D-Optimal Designs using the Uniform Prior

The dependence of the Bayesian criterion on the values of c and d raises a natural question: Is a new design required each time a different prior is placed on β_1 ? Fortunately, a new design is not required provided the change is an alteration of the endpoints of the prior resulting from multiplication by a positive constant. For example, the optimal Bayesian design obtained when $\beta_1 \sim U(-2, -1)$ is the same design that is obtained when $\beta_1 \sim U(-4, -2)$. However, the design obtained from a location shift of, say, -0.5 when $\beta_1 \sim U(-2.5, -1.5)$ is not the same as the optimal design when $\beta_1 \sim U(-2, -1)$. This property will be called ratio constancy because the optimal design remains the same as long as the ratio of the endpoints of the prior on β_1 remain constant. Mathematically speaking, a Bayesian optimal design is ratio constant if the design for $\beta_1 \sim U(c, d)$ is the same as the optimal design when $\beta_1 \sim U(ac, ad)$ where $a > 0$. Of course, this ratio invariance only applies to design points as ECs. Obviously, a change in parameter ranges would cause corresponding change in the natural units of toxicant for the design points.

Ratio constancy of the maximum of (3.3.4) can be shown by using basic algebra techniques. When $\beta_1 \sim U(c, d)$, assume that x_1 and x_2 are the levels of the optimal design which maximize (3.3.4). If $\beta_1 \sim U(ac, ad)$, let x_1^* and x_2^* be the levels of the regressors which maximize (3.4.1)

$$\int_{\beta_1} R(\delta, \beta_1) \pi(\beta_1) d\beta_1 = \frac{n_1 n_2 (x_1 - x_2)^2 (e^{ad(x_1+x_2)} - e^{ac(x_1+x_2)})}{(ad - ac)(x_1 + x_2)} \quad (3.4.1)$$

Multiplying the right hand side of (3.4.1) by $\frac{a^2}{a^2}$ yields

$$\frac{1}{a^2} \frac{n_1 n_2 (ax_1^* - ax_2^*)^2 (e^{ad(x_1^*+x_2^*)} - e^{ac(x_1^*+x_2^*)})}{(d - c)(ax_1^* + ax_2^*)} \quad (3.4.2)$$

In order for the design to be ratio constant, it must be shown that $x_i^* = \frac{x_i}{a}$. By simply grouping ax_i^* together in (3.4.2) it is easy to see that $x_i = ax_i^*$ and $x_i^* = \frac{x_i}{a}$.

The final step of this proof involves showing that the effective concentration of the design points are equivalent for designs with the same ratio. Recall that the value of q_i in the EC_{100q_i} is given by $q_i = e^{\beta_1 x_i}$. When $\beta_1 \sim U(c, d)$ and the maximum value of the Bayes risk occurs at x_i , the value of q_i is given by $e^{\frac{(c+d)}{2} x_i}$. When $\beta_1 \sim U(ac, ad)$ and the maximum value of the Bayes risk occurs at x_i^* , the value of q_i^* is given by $e^{\frac{a(c+d)}{2} x_i^*}$. Since $x_i^* = \frac{x_i}{a}$, performing simple algebraic substitutions shows that the ECs are equivalent. Thus,

$$q_i = e^{\left(\frac{c+d}{2}\right)x_i} = e^{a\left(\frac{c+d}{2}\right)x_i^*} = e^{\left(\frac{c+d}{2}\right)a\left(\frac{x_i}{a}\right)} = e^{\left(\frac{c+d}{2}\right)x_i} = q_i^* \quad (3.4.3)$$

and it is concluded that Bayesian D-optimal designs are ratio constant.

§3.5 Two Level D-Optimal Designs under the Uniform Prior

Two level designs were found using the Nelder-Mead algorithm (Nelder and Mead, 1965). In this case, the D-optimal Bayesian criterion was simple enough that these designs could be confirmed analytically on MAPLE, a mathematics software package (Hardy and Walker, 1995). Table 3.5.1

catalogs these designs by the ratio of the upper and lower endpoints of the prior. The non-control EC is listed in the second row. Allocation percentages for design points are not included in this table because it is easy to see that the criterion (3.3.4) will be maximized when 50% of the observations are at the control and 50% are at an EC specific to the prior placed on the slope parameter. When the ratio is one the optimality criterion is based on a point prior and is thus equivalent to that of Chiacchierini (1996). Also, the design points tend to spread away from each other as the ratio increases or equivalently as the variance in the prior increases.

Table 3.5.1: Two Level D-Optimal Designs under the Uniform Prior

| Ratio | 1 | 1.5 | 2 | 2.5 | 3 | 3.5 | 4 |
|--------------------------|---------------------|---------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| Non-Control Design Point | EC _{13.53} | EC _{12.80} | EC _{11.4} | EC _{9.96} | EC _{8.55} | EC _{7.24} | EC _{6.07} |

§3.5.1 An Example

A simple example in the two level case solidifies the application of these designs in a real life setting. Using the Air Force jet fuel experiment as a backdrop, say that the researcher is interested in studying the effect of the jet fuel on a specific type of plankton present in ocean water. In the absence of toxicant there are typically around 1.3 million organisms in 100 ml of water. The quantity 1.3 million is the expected value at the control or $\lambda_c = e^{\beta_0}$ which implies that $\beta_0 = 14.7$. Based on preliminary observations of the system, the researcher is willing to assume that the EC_{50} falls between (0.2888,1.1552). Translating the interval on the EC_{50} into a uniform prior on β_1 by

the formulas $c = \frac{\ln q_i}{EC_L} = \frac{\ln 0.5}{0.2888}$ and $d = \frac{\ln q_i}{EC_U} = \frac{\ln 0.5}{1.1552}$ gives $\beta_1 \sim U(-2.4, -0.6)$. The mean of this distribution is $\frac{-2.4 + -0.6}{2} = -1.5$.

Now, the optimal design must be found in natural units. Since the ratio of the endpoints is 4, the optimal two level design from Table 3.5.1 is that which places 50% of the observations at the control and 50% of the observations at the $EC_{6.07}$. To obtain the natural units of toxicant (ml) for the non-control design point, the first equation in 3.5.1 is solved for x_1 .

$$\begin{aligned}
q_1 &= e^{\frac{(c+d)}{2}x_1} & (3.5.1) \\
0.0607 &= e^{-1.5x_1} \\
x_1 &= 1.87 \text{ ml}
\end{aligned}$$

So, if 180 experimental units are available, 90 of them will receive no toxicant and 90 of them will receive a dose of toxicant equivalent to 1.87 ml. Later in this work this example will be revisited for a corresponding Bayesian D-optimal design based on the normal prior and an F-optimal design based on the uniform prior.

§3.6 Three Level Bayesian D-Optimal Criterion Under the Uniform Prior

The traditional flaw with two level designs is their failure to provide lack of fit testing. In addition, one might expect them to be less robust to incorrect parameter guesses in comparison to designs with more levels. These familiar problems serve as the usual segue into three level designs. The D-optimality criterion for the three-level design as specified by Chiacchierini is given in (3.6.1).

$$\begin{aligned}
|I| &= n_1 n_2 e^{\beta_0 + \beta_1 x_1} e^{\beta_0 + \beta_1 x_2} (x_1 - x_2)^2 & (3.6.1) \\
&+ n_1 n_3 e^{\beta_0 + \beta_1 x_1} e^{\beta_0 + \beta_1 x_3} (x_1 - x_3)^2 \\
&+ n_2 n_3 e^{\beta_0 + \beta_1 x_2} e^{\beta_0 + \beta_1 x_3} (x_2 - x_3)^2.
\end{aligned}$$

Substituting (3.6.1) into the Bayesian design criterion as $R(d,b)$ yields:

$$\begin{aligned}
\int_{\beta_1} R(\delta, \beta_1) \pi(\beta_1) d\beta_1 &= \frac{n_1 n_2 (x_1 - x_2)^2 (e^{d(x_1+x_2)} - e^{c(x_1+x_2)})}{(d-c)(x_1 + x_2)} & (3.6.2) \\
&+ \frac{n_1 n_3 (x_1 - x_3)^2 (e^{d(x_1+x_3)} - e^{c(x_1+x_3)})}{(d-c)(x_1 + x_3)} \\
&+ \frac{n_2 n_3 (x_2 - x_3)^2 (e^{d(x_2+x_3)} - e^{c(x_2+x_3)})}{(d-c)(x_2 + x_3)}
\end{aligned}$$

(Note that the three level design criterion is consistent with the two level case with respect to its invariance to the prior on β_0 .) While the Nelder-Mead algorithm has been used in the past to

maximize the criterion, it was thought to yield unstable results due to the complicated nature of the function and the number of parameters to be estimated. As a result, a type of Monte Carlo integration that accommodates the most important principle of the Bayesian D-optimal design criterion was pursued. The Monte Carlo method simulates the expected value of the traditional criterion over a distribution by generating a large number of variates from that distribution, evaluating the negative of the Bayes risk function at those values, and then taking the average of the evaluated functions.

In this specific case, β_0 was removed from (3.6.1) to obtain (3.6.3).

$$\begin{aligned} & n_1 n_2 e^{\beta_1 x_1} e^{\beta_1 x_2} (x_1 - x_2)^2 \\ & + n_1 n_3 e^{\beta_1 x_1} e^{\beta_1 x_3} (x_1 - x_3)^2 \\ & + n_2 n_3 e^{\beta_1 x_2} e^{\beta_1 x_3} (x_2 - x_3)^2 \end{aligned} \tag{3.6.3}$$

This expression was evaluated at 1000 random variates generated from a $U(c,d)$ distribution over a fine grid of each combination of the six components that make up the design $(p_1, p_2, p_3, q_1, q_2, q_3)$ where $n_i = p_i N$. These 1000 values were averaged at each combination of the design points and allocation percentages to obtain the approximate integral. The values of $(p_1, p_2, p_3, q_1, q_2, q_3)$ corresponding to the maximum function value were then selected as the optimal design. Note that this process simulates finding the maximum of (3.6.2). Consistent with the cases thus far, the optimal three level design degenerated to two levels. Hence restrictions were placed on the designs to yield useable three level designs.

§3.7 Type I Three Level Bayesian D-Optimal Designs under the Uniform Prior

For the restricted designs, one level was set at the control and the other two levels were required to be symmetric about the center of the region of interest. In addition the percentage of the design points allocated at the two non-control levels was specified to be equal. These designs were also developed specifically with equal allocation at the non-control points because researchers often desire designs with this property. They will be referred to as Type I three level designs.

Chiacchierini (1996) used these same restrictions for her development of three level designs in the non-Bayesian case.

In the development of Type I designs, three regions of interest are considered: $[EC_5, EC_{80}]$, $[EC_{10}, EC_{80}]$, and $[EC_{20}, EC_{80}]$. Tables A1-A3 in Appendix A contain these designs and their efficiencies. The formula for a design efficiency in the D-optimal case is

$$E = \left(\frac{|\mathbf{I}| \text{ for the design of interest}}{|\mathbf{I}| \text{ for the optimal design}} \right)^{\frac{1}{p}} \quad (3.7.1)$$

where p is the number of parameters in the model. Efficiencies are calculated with respect to the optimal restricted three level designs on the corresponding region of interest where the parameters are known. The efficiencies of these Bayesian designs are cataloged by the ratio of $\frac{c}{d}$. A graphical look at these efficiencies is a quick and easy way to gain insight into the performance of these designs. In the graph below, efficiencies are plotted against this ratio for the designs on each region of interest. Designs on different regions are denoted by the lower cutoff for their respective regions of operability.

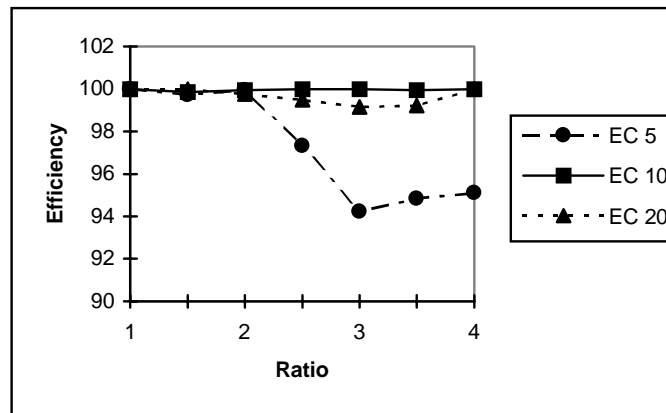


Figure 3.7.1 Efficiencies vs. Ratio for Type I D-Optimal Uniform Designs.

Several comments are in order about these designs. From the graph, one can see that the overall efficiencies of these designs are extremely good, the lowest one being approximately 94%.

Additionally, a glance through the design tables in Appendix A show that these designs are fairly insensitive to the prior on β_1 , particularly for the restricted regions $[EC_{10}, EC_{80}]$, and $[EC_{20}, EC_{80}]$. In fact, they are virtually the same designs as detailed in Chiacchierini (1996). The designs on the $[EC_5, EC_{80}]$ are more sensitive to the prior as evidenced by their differentiation into two groups. There is a group of similar designs for ratios less than or equal to two and a group of similar designs for ratios greater than two. This can be seen in the dip in the graph of the efficiencies for the region $[EC_5, EC_{80}]$ around the ratio of two. It is at this point that the more variable prior on β_1 causes the design points to spread further from the optimal design. Designs on the other two regions do not experience a similar decline in efficiency with the wider prior because the lower bound of the region is particularly close to the same optimal level for Chiacchierini's symmetric designs (1996).

§3.8 Robustness Properties of Type I Designs under the Uniform Prior

The development of these Bayesian designs based on varying amounts of information about the parameter β_1 implies that one would want to consider the robustness properties of these designs with respect to parameter misspecification. One would think that designs based on more variable priors would be more robust in the presence of misspecification than the most efficient two level design. While a three level design will never be considered more efficient than a two level design under known parameters (Chiacchierini, 1996), it is desirable to find three level designs that are efficient in the presence of parameter misspecification. Restricting the designs to equal allocation at the non-control points may be advantageous from a practical point of view but it severely impacts efficiency when parameters are known. In turn, the three level designs are often unable to “catch-up” under the presence of misspecification. Increasing these efficiencies in both situations requires only a simple modification of the restrictions on the existing three level criterion.

In general, the modification makes the three level designs more similar to the two level case and thus more efficient and more robust in many cases. This is accomplished by restricting the design so that experimental units are allocated equally to the control and the lower end EC. Recall that the two level D-optimal design for a point prior places 50% of the observations at the $EC_{13.53}$ and 50% at the control. By forcing equal sample sizes at the non-control levels, as in the Type I design, the three level design is forced further away from optimality. This departure from optimality

results from the criterion's tendency to distribute points equally at all three ECs. By placing a large percentage of design points in between the control and the lower end EC, the design is pushed further away from the optimum. For that reason, a second type of three level Bayesian designs was generated.

§3.9 Type II Three Level Bayesian D-Optimal Designs under the Uniform Prior

The first two restrictions placed on the Type II designs are the same as the restrictions for the Type I designs. One design point was required to be the control and the non-control points were restricted to be symmetric around the region of operability. However, the third constraint regarding sample size was altered to provide for equal allocation at the control and the lower end EC. This makes the three level design more similar to the two level design. Of course, this design would degenerate to the two level design so the criterion was forced to place at least 10% at the other non-control EC. The designs were generated on the same three regions of operability as the Type I designs: $[EC_5, EC_{80}]$, $[EC_{10}, EC_{80}]$, and $[EC_{20}, EC_{80}]$. The Monte Carlo method was used to approximate the expected value of the criterion in (3.6.3). The designs which maximized the criterion for their respective regions and priors are tabled in A4-A6 in Appendix A. Graphs of the efficiencies of the Type II designs against the optimal three level designs in Chiacchierini are displayed in Figure 3.9.1. It shows that Type II Bayesian D-optimal designs outperform Chiacchierini's designs in every situation.

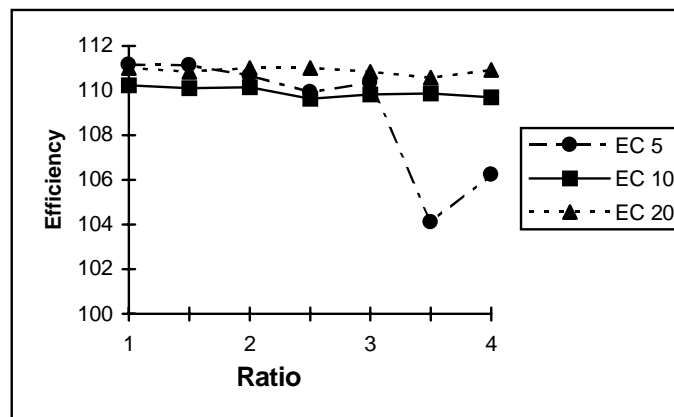


Figure 3.9.1 Efficiencies vs. Ratio for Type II D-Optimal Uniform Designs.

§3.10 Robustness Study for Type II Three Level Designs under the Uniform Prior

Finally, the robustness issue must be addressed. The following study demonstrates the performance of these designs in the presence of misspecification of the slope parameter. In these designs the misspecification of the prior on β_1 is a direct result of misspecification of the range on the EC from the biologist. In the two level design example, the researcher specified an interval on the EC_{50} . In the robustness study, it is assumed that the biologist made an error and the EC_{50} is not the true EC for the given range. Rather, it covers another EC such as the EC_{30} or the EC_{55} . As one might imagine, the prior placed on β_1 and the value of β_1 derived from a range on the EC_{30} or EC_{55} are both quite different from the same quantities when they are developed under the assumption that the interval on the EC_{50} is correct. An example makes this more clear.

Suppose that the biologist states the EC_{50} falls between 0.2888 and 1.1552. In the two level Bayesian design example, this prior is translated to a prior on β_1 where $\beta_1 \sim U(-2.4, -0.6)$ and the mean of this distribution is $\frac{-2.4 + -0.6}{2} = -1.5$. Contrast this scenario with that where the range given for the EC_{50} is really the true range for the EC_{30} . So instead of EC_{50} falling between 0.2888 and 1.1552, the EC_{30} is contained in (0.2888, 1.1552). This changes the prior on β_1 to $\beta_1 \sim U(-4.169, -1.042)$ and the true $\beta_1 = \frac{-4.169 + -1.042}{2} = -2.605$. The curve based on the true β_1 is then a much steeper curve than the one based on the estimate of $\beta_1 = -1.5$. Thus the design points, in natural units, based on the assumed estimate of $\beta_1 = -1.5$ are decidedly off target. The robustness study shows how Type II three level designs perform under this type of parameter misspecification.

For brevity, only three designs were considered in the robustness study. These include the three level designs based on the point prior and priors with ratios of 2.5 and 4. The prior on the EC_{50} was then assumed to be a prior on one of the following: EC_{30} , EC_{40} , EC_{45} , EC_{55} , EC_{60} , and EC_{70} . For each of these six possibilities the true β_1 was computed. The mean of the prior on β_1 , or the assumed β_1 , in each of the designs studied is -1.5. The robustness study is tabled in A7-A9 in Appendix A. The true values of β_1 and the true ECs are shown in the tables. The efficiencies of the three level designs versus the amount of misspecification is graphed for each of the three designs in

Figure 3.10.1. Efficiencies are calculated with respect to Chiacchierini's optimal *two level design*. For designs based on more variable priors, efficiencies increase as the amount of misspecification from the EC_{50} increases and vice versa as the amount of misspecification decreases from the EC_{50} . The designs based on the three priors studied increase in efficiency as the EC misspecification increases. While some of the efficiencies are not favorable for the Type II designs, the efficiencies with respect to the two level design are still greater than 100% in some cases. Two conclusions can be drawn from this graph. Type II designs appear to be more robust when the range on the EC_{50} really covers a higher EC. In the case where the range on the EC_{50} contains a lower EC, the two level design seems more robust.

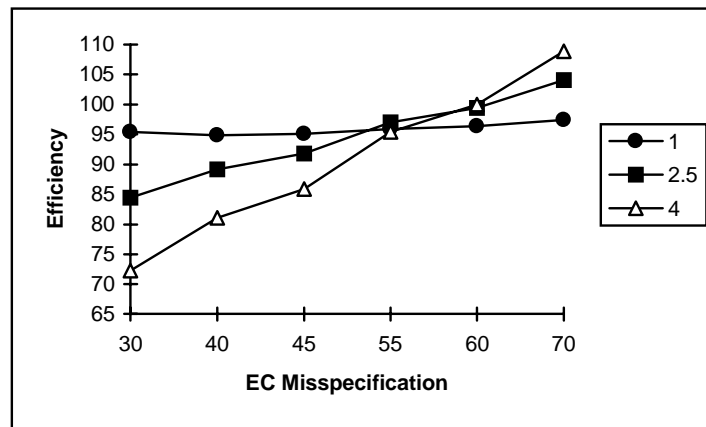


Figure 3.10.1 Two Level Efficiencies of Type II D-Optimal Uniform Designs under Misspecification.

A second graph is included to characterize the efficiencies of the Type II designs with respect to Type I designs. From Figure 3.10.2, one can conclude that Type II designs are superior to

Type I designs in every case and that they are particularly efficient when misspecification occurs with respect to the lower end. These efficiencies are also tabled in Appendix A.

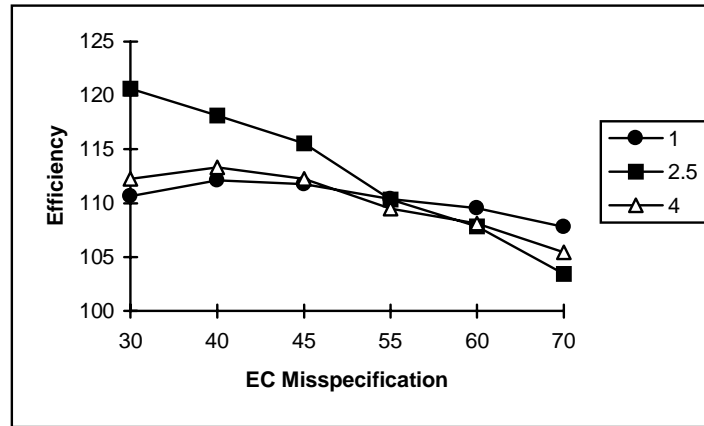


Figure 3.10.2 Three Level Efficiencies of Type II D-Optimal Uniform Designs under Misspecification.

§3.11 The Normal Prior on β_1 in the Exponential Model

While the uniform prior on β_1 provides many nice Bayesian designs, the assumption that β_1 takes on all values in a range with equal probability many not be valid. Instead, the biologist may be inclined to place more weight on those values in the center of the range. Hence, the introduction of Bayesian designs based on a normal prior on β_1 is appropriate.

As in the uniform case, the normal prior on β_1 originates from a range on a particular EC denoted by (EC_L, EC_U) . Using the formulas $\beta_L = \frac{\ln q_i}{EC_L}$ and $\beta_U = \frac{\ln q_i}{EC_U}$, this interval is translated

to a range on β_1 . Now, these values can be used to find the mean and standard deviation of a normal distribution on β_1 . The average of β_L and β_U is taken to be the mean of the normal distribution, μ_{β_1} . The standard deviation, σ_{β_1} , is found by equating β_L and β_U to the z-scores of $z_{\alpha/2}$ and $-z_{\alpha/2}$

where $\frac{\alpha}{2}$ represents the right tail area in the normal distribution and solving for the value of σ_{β_1} .

The value of α is chosen at the discretion of the researcher depending on the certainty of his or her original estimates of the ECs. All designs in this work are based on $\alpha = 0.05$.

§3.12 Two Level Bayesian D-Optimality Criterion under the Normal Prior

Substituting the D-optimality criterion (3.3.1) and the normal prior into the Bayesian design criterion yields

$$\begin{aligned} \int_{\beta_1} R(\delta, \beta_1) \pi(\beta_1) &= \int_{-\infty}^{\infty} n_1 n_2 e^{\beta_1 x_1} e^{\beta_1 x_2} (x_1 - x_2)^2 \frac{1}{\sqrt{2\pi\sigma_{\beta_1}}} e^{-\frac{(\beta_1 - \mu_{\beta_1})^2}{2\sigma_{\beta_1}^2}} d\beta_1 \\ &\propto n_1 n_2 e^{\mu_{\beta_1}(x_1+x_2) + (x_1+x_2)\sigma_{\beta_1}^2}. \end{aligned} \quad (3.12.1)$$

Like the uniform case factors involving β_0 , are eliminated from the criterion because they are not directly tied to the choice of the design points.

Due to the complicated nature of the function, Monte Carlo integration was pursued instead of numerical optimization via Nelder Mead for both two and three level designs. The algorithm for Monte Carlo integration in Section 3.6 was used for this case as well. These designs are discussed in Sections 3.14-3.17.

§3.13 Sigma Constancy Property of Designs based on the Normal Prior

While designs based on the uniform prior could be cataloged by the ratio of the endpoints of the prior on β_1 , designs based on the normal prior can be cataloged by the value of σ_{β_1} . The parameter μ_{β_1} has no impact on the design points chosen through the maximization process because its only involvement in the function comes in the form of ECs. When μ_{β_1} appears in (3.12.1), it appears as $q_i = e^{\mu_{\beta_1} x_i}$ as in EC_{100q} . For this reason, the same value of q and hence the same EC_{100q} will maximize (3.12.1). While the values of the toxicant in natural variables will vary with different values of μ_{β_1} , the EC_{100q} which make up the optimal design will not change. However, one can see that the value of σ_{β_1} will affect the ECs chosen as design points. Simulation confirms that the same values of σ_{β_1} produce the same design regardless of μ_{β_1} . Hence, these designs can be categorized as sigma constant.

§3.14 Two Level Designs under the Normal Prior

The two level designs based on the normal prior are similar to those based on the uniform prior. The Bayesian D-optimality criterion is maximized when 50% of the observations are placed at the control and the remaining 50% of the observations are placed at a non-control level specific to the value of σ_{β_1} . In the designs listed in Table 3.14.1, there is a direct correspondence between the value of σ_{β_1} and the ratio of the endpoints of the uniform prior. The lower and upper bounds on the 95% confidence interval are the same as the bounds for the uniform prior. The non-control levels tend to spread out as sigma increases. However, the levels do not spread out as far as they did in the uniform case for corresponding designs, especially for those designs with more variable priors.

Table 3.14.1 Two Level D-Optimal Designs under the Normal Prior

| σ_{β_1} | 0.00005 | 0.1500 | 0.2500 | 0.3214 | 0.3750 | 0.416 | 0.4500 |
|--------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|--------------------|--------------------|
| Non-Control Design Point | EC _{13.53} | EC _{12.58} | EC _{11.85} | EC _{11.64} | EC _{10.62} | EC _{9.80} | EC _{9.37} |

§3.14.1 Two Level Design Example Revisited

An example of a two level Bayesian D-optimal design was presented in Section 3.5.1 based on the toxicity studies that the United States Air Force was performing. In this example, a design was examined where EC_{50} fell between (0.2888,1.1552). For the normal example the EC_{50} will be contained in the same range. However, once the range is transformed to an interval on β_1 , it corresponds to a 95% normal probability interval on β_1 with endpoints of -2.4 and -0.6 and a mean of -1.5. Using these quantities, the value of $\sigma_{\beta_1} = 0.45$. The two level design which corresponds to this value of σ_{β_1} places 50% of the observations at the $EC_{9.37}$ and 50% at the control. Solving for these points in natural units gives

$$q_1 = e^{\frac{(c+d)}{2}x_1} \tag{3.12.1}$$

$$0.0937 = e^{-1.5x_1}$$

$$x_1 = 1.58 \text{ ml.}$$

If 180 experimental units are available, this design indicates that 90 of them should be treated with 1.58ml of fuel and 90 of them with no fuel. Note that the amount of toxicant at the non-control point decreases by 18.4%, 1.87ml to 1.58ml, from the comparable design based on the uniform prior. This indicates that two level Bayesian designs are indeed sensitive to the type of prior placed on the slope parameter.

§3.15 Type I Three Level Designs under the Normal Prior

The criterion for the three level design is found via the same methods used for designs based on a uniform prior. Of course, an initial attempt was made to use this criterion to determine unrestricted three level designs based on the normal prior. However, this attempt was unsuccessful since the optimal designs found with the unrestricted three level criterion only had two levels. In order to obtain practical three level designs, Type I restrictions were placed on the designs. One design point was specified to be the control, the two remaining points were restricted to be symmetric around the center of the region and the percentage of allocation at the non-control points was forced to be equal. The efficiencies for these designs versus their corresponding value of σ_{β_1} are graphed for each region in Figure 3.15.1. The actual designs are tabled in Appendix A.

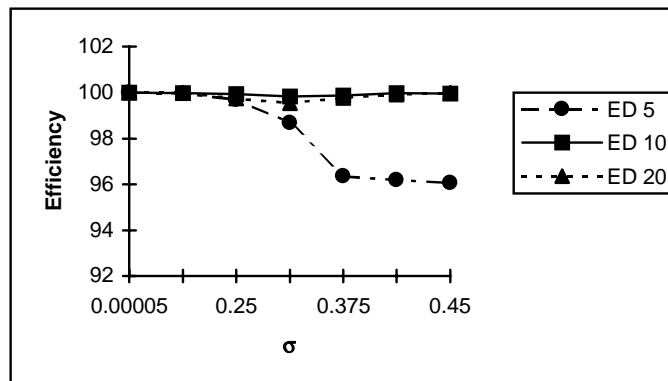


Figure 3.15.1 Efficiencies vs. Sigma for Type I D-Optimal Designs under the Normal Prior.

The efficiency patterns of these designs are very much like those seen in Figure 3.7.2 for the uniform case. Robustness to the prior is a recurring theme in three level Bayesian designs on the $[EC_{10}, EC_{80}]$ and $[EC_{20}, EC_{80}]$. As the prior becomes more variable on these regions, the designs remain similar in both the design points and the amount allocated to those points. The designs on

the region defined by the $[EC_5, EC_{80}]$ show the same differentiation into groups as their uniform counterparts. On this region, the efficiency drops noticeably when $\sigma_{\beta_1} = 0.3214$. Also, it appears that a general comment can be made about the robustness of the Type I three level Bayesian designs to the type of prior- uniform or normal. Designs based on the normal prior are comparable to corresponding designs based on the uniform prior in terms of the design points and the percentage of allocation to these points. This is evident through the graphs and tables for these designs.

§3.16 Type II Three Level Bayesian D-Optimal Designs

While the Type I designs are desirable for the researcher because of equal allocation at the non-control points, they again fail the robustness test in comparison to the Type II designs. Type II designs require one level to be the control and the two non-control levels to be symmetric around the center of the region of operability. The sample size restrictions differ from the Type I designs in that the control and the lower end EC are required to have the same sample size. Type II designs are tabled in Appendix A. Their efficiencies with respect to Chiacchierini's Type I designs are graphed in Figure 3.16.1. Notice that all of the designs exhibit a superior performance when compared to the status quo.

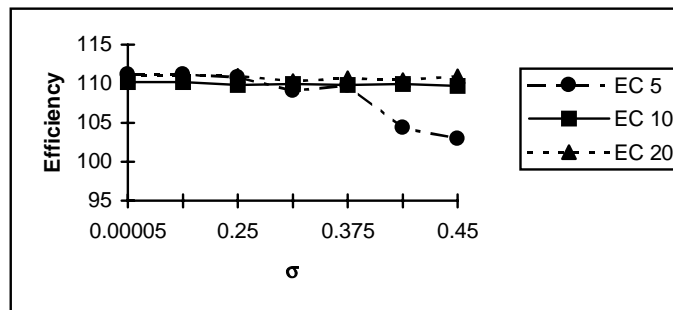


Figure 3.16.1 Efficiencies vs. Sigma for Type II D-Optimal Designs under the Normal Prior.

§3.17 Robustness Study for Type II Designs

The robustness properties of the Type II designs were explored in terms of their efficiencies against the two level design in the face of parameter misspecification. The robustness study includes three designs on the (EC_5, EC_{80}) with standard deviations of 0.00005, 0.3214, and 0.45. The true values of β_1 were obtained as a result of the prior on the EC_{50} actually being a prior on one of the

following: EC_{30} , EC_{40} , EC_{45} , EC_{55} , EC_{60} , EC_{70} . The results of the study are cataloged in Tables A16-A18 in Appendix A. Graphs of the efficiencies against the two level design are shown in Figure 3.17.1. One line is plotted for each of the three designs. The lines are designated by the corresponding value of σ_{β_1} .

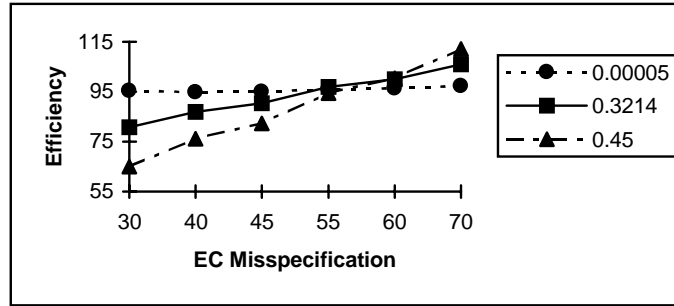


Figure 3.17.1 Two Level Efficiencies of Type II D-Optimal Normal Designs under Misspecification.

Like the uniform case, the efficiencies listed in Figure 3.17.1 indicate that Type II three level designs are more efficient if the initial range for the EC is misspecified upward. Also, we see that the two level designs perform quite well under upper end misspecification. Figure 3.17.2 shows the Type II designs are the better than the Type I in almost all situations.

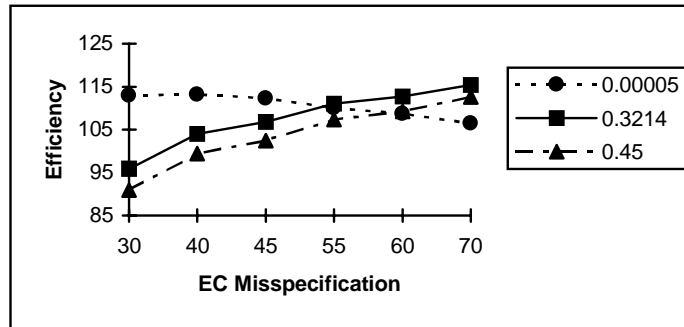


Figure 3.17.2 Three Level Efficiencies of Type II D-Optimal Normal Designs under Misspecification.