Chapter 9
Modified Robust Design Criteria for Poisson Mixed Models

Hongyan Jiang and Rongxian Yue

Abstract The maximin D-optimal design (MMD-optimal design) and hypercube design (HCD-optimal design) are two robust designs which overcome the problem of design dependence on the unknown parameters. This article considers the robust designs for Poisson mixed models. Given the prior knowledge of the fixed effects parameters, a modification of the two robust design criteria is proposed by applying the number-theoretic methods. The simulated annealing algorithm is used to find the optimal exact designs. The results show that the modified optimal designs perform better in the relative *D*-efficiency and programming time.

9.1 Introduction

In the fields of optimal experimental design, the Fisher information matrix plays an important role. For nonlinear models or generalized linear models, the Fisher information matrix depends on the unknown values of the parameters, which means that the optimal design will depend on the parameters. Researchers can fix the value based on their knowledge, or just guess, then the design will be locally optimal.

Robust design criterion is a good choice to overcome the problem of dependence of a design on the unknown parameters, such as the maximin criterion and Bayesian criterion [4]. The Bayesian approach maximizes the expected Shannon information considering the prior information about the parameters of the model, while the maximin approach optimises over a specific domain of parameter values by maximizing the minimal value of a measure of the information matrix, in which the parameters are assumed to belong to a known domain, without any hypothesis on their underlying distribution [2].

H. Jiang
Huaiyin Institute of Technology, Jiangsu 223003, China
e-mail: hyitjhy@163.com

R. Yue (✉)
College of Mathematics and Science, Shanghai Normal University, Shanghai 200234, China
e-mail: yue2@shnu.edu.cn

J. Fan and J. Pan (eds.), *Contemporary Experimental Design, Multivariate Analysis and Data Mining*,
https://doi.org/10.1007/978-3-030-46161-4_9

Aside from classical robust design criteria, the product design criterion, first suggested by Atkinson and Cox [1], maximized the product of the determinants of Fisher information matrices of the models of interest, scaled to the number of parameters in each model. McGree et al. [11] applied the product design criterion to optimise the product of the normalised determinants of Fisher information over eight different mixed effects bio-impedance models, which was combined by the 2.5th and 97.5th percentiles of all four fixed effect parameters in the model. Foo and Duffull [9] proposed a hypercube D-optimality (HCD) criterion and a hypercube maximin D-optimality (HCMMD) criterion, by setting the domain Θ_{HC} of the fixed effect parameters as various combinations of the 2.5th and 97.5th percentiles from the known prior distribution of them in nonlinear mixed models. The HCD method is a particular case of the product design criterion, and the result shows that this method performs better at some combination of the extrema values of the parameters. What's more, a 100-fold improvement in the speed of this method compared to the Bayesian optimal design is particularly attractive.

However, the percentiles of the prior distribution do not scatter as 'uniform' as possible, and the underlying assumption of the HCD and HCMMD is that the efficiency of any locally D-optimal design of the 97.5% percentiles is more or as efficient to design of the parameter values located within the 97.5% interval [9]. We want to generate a set of the parameter values which are uniformly scattered in a given multi-dimensional prior distribution. Number-theoretic methods (NTMs) are used in experimental design by Fang and Wang [5]. The set of the representative points (RPs) based on NTMs is uniformly scattered under the notation of discrepancy. The aim of this paper is to provide a robust method of obtaining optimal designs based on the RPs. In what follows, given the prior distribution of the fixed effect parameters, a D-optimality criterion based on the set of RPs, denoted by RPD-optimality criterion, and a maximin optimality criterion based on the set of RPs, denoted by RPMMD, are proposed.

The rest of the paper is organized as follows. The Poisson mixed models are introduced in Sect. 9.2. Section 9.3 gives a brief review on the existing criteria, and presents a modification of the robust criteria by using the NTMs. Section 9.4 evaluates the new robust criteria via an one-variable first-order and second-order Poisson mixed models by comparing among several designs. Section 9.5 is the conclusion of the paper.

9.2 The Poisson Mixed Model

In this section, a Poisson mixed model is introduced, and the quasi-likelihood method [12, 13, 16] is applied to Poisson mixed model to obtain the quasi-information matrix.

9.2.1 Poisson Mixed Models

Suppose there are N independent individuals taken part in an experiment, and the responses y_{ij} at the experimental settings x_{ij} of an explanatory variable x for individual i follows a Poisson distribution, conditioned on an r-dimensional random effects vector b_i [12, 13]. It is assumed that y_{ij}'s are related to the fixed and random effects via a log link, that is $\log(\lambda_{ij}) = f_{ij}^T\beta + z_{ij}^T b_i$, and

$$p(y_{ij}|b_i) = \frac{\lambda_{ij}^{y_{ij}}}{y_{ij}!}\exp(-\lambda_{ij}), \quad y_{ij} = 0, 1, 2, \cdots, \quad i = 1, 2, ..., N, \quad j = 1, 2, \cdots, m_i, \tag{9.1}$$

where $p(y_{ij}|b_i)$ denotes the conditional probability density function of y_{ij} given b_i. Moreover, given the individual random effects b_i, the observations y_{ij} are assumed to be conditionally independent. The $p \times 1$ vector f_{ij} is the design vector of the explanatory variable at the jth measurement for individual i, β is the corresponding $p \times 1$ vector of unknown fixed effect parameters, z_{ij} is the $r \times 1$ $(r \leq p)$ design vector for the random effects which is usually a subset of vector f_{ij}, and $b_i, i = 1, 2, \ldots, N$ is the corresponding $r \times 1$ vector of unknown random effects which are drawn independently from a multivariate normal distribution with mean zero and covariance matrix G.

Let the vector $y_i = (y_{i1}, \cdots, y_{im_i})^T$ be the count responses of individual i, and $y = (y_1^T, y_2^T, \cdots, y_N^T)^T$ be the response vector of the experiment for the N individuals.

9.2.2 Fisher Information Matrix of the Model

Our interest lies in measuring the responses under a reasonable experimental design to estimate the fixed effect parameter β as accurately as possible. For simplicity we will assume throughout that the covariance matrix G of b_i is known. Note that the covariance matrix G need not be of full rank, which allows for some or most of the parameters to be fixed across the individuals.

The likelihood function of β is

$$L(\beta) = \prod_{i=1}^{N}\int\prod_{j=1}^{m_i} p(y_{ij}|b_i)p(b_i)db_i, \tag{9.2}$$

where $p(b_i)$ is the probability density function of b_i. The maximum likelihood estimator of β cannot be written down in closed form due to the random effects in model (9.1). As mentioned in [12–14, 16], quasi-likelihood method is employed to construct the quasi-likelihood function $QL(\beta; y)$. See [8, 10, 16] for details. The quasi-information matrix for the experiment is

$$M(\beta) = D^T V^{-1}(\mu(\beta))D = \sum_{i=1}^{N} M_i(\beta), \tag{9.3}$$

where $\mu(\beta)$ is the marginal mean of y, $V(\mu(\beta))$ is the marginal covariance matrix of y, $D = \partial\mu(\beta)/\partial\beta$, and $M_i(\beta)$ is the quasi-information matrix of individual i.

According to the technique of variance correction in [14], we define a variance correction term

$$c(z_{ij}, z_{ij'}) = \exp(z_{ij}^T G z_{ij'}) - 1,$$

and let $C_i = (c(z_{ij}, z_{ij'}))$ be the $m_i \times m_i$ matrix of the correction terms. Then the quasi-information matrix of individual i is given by

$$M_i(\beta) = F_i^T A_i^T (A_i + A_i C_i A_i)^{-1} A_i F_i = F_i^T (A_i^{-1} + C_i)^{-1} F_i, \tag{9.4}$$

where A_i is a diagnose matrix with the individual mean vector $E(Y_i)$ on its diagonal. Note that

$$D_i = \frac{\partial\mu_i(\beta)}{\partial\beta} = A_i F_i,$$

where $F_i^T = (f_{i1}, \cdots, f_{im_i})$ is the design matrix of individual i.

In what follows we mainly consider the one-variable first-order Poisson mixed model

$$\lambda_{ij} = \exp(\beta_0 + b_{i0} + (\beta_1 + b_{i1})x_j), \tag{9.5}$$

and the one-variable second-order Poisson mixed model

$$\lambda_{ij} = \exp(\beta_0 + b_{i0} + (\beta_1 + b_{i1})x_j + (\beta_2 + b_{i2})x_j^2). \tag{9.6}$$

In these models the design vectors for the fixed effects and the random effects are equal, i.e., $f_{ij} = z_{ij}$ in model (9.1).

9.3 Robust Optimal Designs

9.3.1 Locally D-Optimal Designs

In most practical situations, exact design with a given total number of design points is required. The objective of this paper is to determine an optimal m-exact design of the following form

$$\xi_m = \begin{Bmatrix} x_1 & x_2 & \cdots & x_s \\ n_1 & n_2 & \cdots & n_s \end{Bmatrix},$$

where x_k, $k = 1, 2, \cdots, s$, are the s different settings for each individual, and n_k denotes the corresponding repetition times of observations at x_k, $k = 1, 2, \cdots, s$, and $\sum_{k=1}^{s} n_k = m$. The individual design with fixed m is considered, which is reasonable in practice. Each exact design can be considered as a design measure over the design region, which can be written as a probability measure with supports x_k's:

$$\xi = \begin{Bmatrix} x_1 & x_2 & \cdots & x_s \\ p_1 & p_2 & \cdots & p_s \end{Bmatrix}, \quad p_k = \frac{n_k}{m}, \quad \sum_{k=1}^{s} p_k = 1.$$

A design ξ that makes the estimation of the unknown parameters in a model, β, as effectively as possible, dominates over all other designs in the set of all design measures Ξ in the Löwner sense is called Löwner optimal. However, it is very difficult to find the Löwner optimal design ξ, in general. A popular way is to specify an optimality criterion, which is defined as a real-valued function of the information matrix $M(\xi; \beta)$ of the model. The most commonly used function is logarithm of its determinant $\log |M(\xi; \beta)|$ and the corresponding optimality is known as D-optimality. A design ξ is called a locally D-optimal in the Poisson mixed model (9.1) if for a given nominal value of β, it maximizes $\log |M(\xi; \beta)|$, i.e.,

$$\xi^D = \arg\max_{\xi} \log |\mathbf{M}(\beta)|. \tag{9.7}$$

It is known that a D-optimal design ξ^D minimizes the content of the confidence region of β and so minimizes the volume of the ellipsoid [2]. Note that the information matrix $M(\xi; \beta)$ for a general model usually depends on the parameters β, and then the design ξ^D is called locally D-optimal. In Sect. 9.4, the locally D-optimal designs for the Poisson mixed models in (9.5) and (9.6) are calculated at the prior means of β, respectively.

Niaparast and Schwabe [14] provides an equivalence theorem for checking the optimality for a given candidate design for the Poisson mixed models. The D-efficiency of an arbitrary design ξ compared to the D-optimal design ξ^D is defined as [2]

$$D_{\text{eff}} = \left(\frac{|M(\xi; \beta)|}{|M(\xi^D; \beta)|} \right)^{\frac{1}{p}}, \tag{9.8}$$

where, p is the number of parameters for the fixed effects of the model.

A Bayesian D-optimal design, ξ^{BD}, helps to overcome the problem of design dependence on the unknown parameters, is defined as follows:

$$\xi^{BD} = \arg\max_{\xi} \int_{\beta} \log |M(\xi; \beta)| \eta(\beta) d\beta, \tag{9.9}$$

where $\eta(\beta)$ is a chosen prior distribution of β. The integration here will be calculated numerically by quasi-Monte Carlo (QMC) methods. It is known that the QMC methods for multi-dimensional numerical integration are much more efficient than traditional Monte Carlo methods [5].

9.3.2 RPD-and RPMMD-Optimalities

Foo and Duffull [9] proposed a hypercube design criterion termed HCD-optimality, which is a specific case of product optimality, where component models are formed by the same structure model but with sets of parameter values taken at the 2.5th and 97.5th percentiles values of the prior distribution of β. A maximin design criterion was also considered in [9] by setting the domain of parameters as Θ_{HC} composed of all the combinations of the 2.5th and 97.5th percentile values, which is called HCMMD-optimality. The HCD-optimal design is defined by

$$\xi^{HCD} = \arg\max_{\xi} \sum_{\beta \in \Theta_{HC}} \log |M(\xi; \beta)|, \tag{9.10}$$

and the HCMMD-optimal design is defined by

$$\xi^{HCMMD} = \arg\max_{\xi} \min_{\beta \in \Theta_{HC}} \log |M(\xi; \beta)|. \tag{9.11}$$

The method in [9] is attractive for its short operating time and acceptable effective at some nominal parameter values. The maximin optimal designs [4, 7, 17] are particularly attractive since an appropriate range for the unknown parameters is only required to specify. The major problem is that the maximin optimality criterion is not differentiable and the equivalence theorem is elusive.

Note that the set of percentiles may not represent as much information of a Multivariate distribution as possible. We now consider the use of RPs of the prior distribution by NTMs. Fang and Wang [5] introduced two kinds of RPs based on the F-discrepancy criterion and MSE criterion, respectively. Under the F-discrepancy criterion, there exists a set of optimal RPs for a given continuous univariate distribution by directly using the inverse transformation method. For the multivariate distributions with independent components, their RPs may also be obtained by using the inverse transformation method. For the multivariate distributions with dependence structures, Fang and Wang [5] proposed the NTSR algorithm to generate their RPs, which can be implemented to obtain the RPs of the spherically symmetric distribution, multivariate l_1-norm distribution, Liouville distribution, and so on. Zhou and Wang [18] considered the RPs of Student's t_n distribution for minimizing the MSE criterion. Very recently, Zhou and Fang [19] proposed a new criterion, termed FM-criterion, to choose n RPs of a given distribution, which minimize the L_2-norm of the difference between the empirical distribution and the given distribution under the constraint that the first $n-1$ sample moments equal the population moments. The empirical study in [19] shows that the RPs under the FM-criterion are better than other types of RPs. It is known that finding RPs under the MSE criterion is more difficult, but more appropriate in the case of small sample size.

In what follows, by $\Theta_{n\text{-}RP}$ we denote a set of n RPs generated by the inverse transformation method under F-discrepancy criterion from a prior distribution of β with independent components. We define two robust design criteria to against the uncertainty of the fixed effects in the mixed model (9.1) by using the RPs in $\Theta_{n\text{-}RP}$, and compare them with the existing criteria in (9.10) and (9.11).

A design is called RPD-optimal if it maximizes

$$\Phi_{RPD}(\xi) = \sum_{\beta \in \Theta_{n\text{-}RP}} \log |M(\xi; \beta)|, \tag{9.12}$$

and a design is called RPMMD-optimal if it maximizes

$$\Phi_{RPMMD}(\xi) = \min_{\beta \in \Theta_{n\text{-}RP}} \log |M(\xi; \beta)|. \tag{9.13}$$

9.4 Numerical Studies

In this section we present Numerical studies for the RPD-and RPMMD-optimal designs for the first-order model in (9.5) with three different covariance structures of the random effects, and the second-order model in (9.6) with a diagonal covariance matrix of the random effects, respectively. The design region is taken as $[c, 1]$ with $c = 0.01, 0.2, 0.4$, respectively, as used in [15].

We assume that β has a continuous multivariate prior distribution $H(\beta)$ with independent components. i.e., $H(\beta) = H(\beta_1, \cdots, \beta_p) = \prod_{i=1}^{p} H_i(\beta_i)$, where $H_i(\beta_i)$ $(i = 1, \cdots, p)$ are the marginnal distribution functions of β. We use the NTMs as demonstrated in [5] to find the set of RPs of the prior distribution. Letting $\{c_k = (c_{k1}, \cdots, c_{kp}), k = 1, \cdots, n\}$ is a set of n points which are uniformly scattered in the unit cube $C^s = [0, 1]^s$, e.g., a good lattice points (glp) set, then the set $\Theta_{n\text{-}RP}$ is obtained by using the inverse transformation method, i.e., $\Theta_{n\text{-}RP} = \{\beta_k = (H_1^{-1}(c_{k1}), \cdots, H_p^{-1}(c_{kp})), k = 1, \cdots, n\}$.

To find the optimal m-exact designs that maximize the criteria defined in last section, we use the simulated annealing (SA) algorithm. In our computation for $m = 8, 12, 24$, the initial temperature in the SA algorithm is taken as $T_0 = 10^6$, and the temperature reduction factor is 0.9. It is known that the SA algorithm allows the search patterns to move away from a path of strict descent, migrates through a sequence of local extremum in search of the global solution, and recognizes when the global extremum has been located [3, 6, 9].

It must be noted that the Bayesian D-optimality criterion (9.9) requires a complicated integration over the prior distribution. The computation of the Bayesian optimal designs involves two steps: (i) computation of criterion for a given design, and (ii) finding an optimal design by maximization of the criterion value. To compute the criterion (9.9) for a given design, we use the NTMs which is more efficient than Monte Carlo methods to obtain a good approximation of integration.

9.4.1 Designs for the First-Order Poisson Mixed Model

For the first-order Poisson mixed model given in (9.5), we consider the following three kinds of covariance matrices G of random effects $b = (b_0, b_1)^T$:

$$G_1 = \begin{pmatrix} 0.5 & 0 \\ 0 & 0 \end{pmatrix}, \quad G_2 = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}, \quad G_3 = \begin{pmatrix} 0.5 & 0.25 \\ 0.25 & 0.5 \end{pmatrix}.$$

9.4.1.1 The Case of Normal Prior Distributions

Assume that the prior distribution of $\beta = (\beta_0, \beta_1)^T$ is a normal distribution with mean $\bar{\beta} = (\bar{\beta}_0, \bar{\beta}_1)^T = (1, -3)^T$ and an identity covariance matrix I_2.

In order to compare with the design criterion using percentile points in Θ_{HC} which contains 4 values of β, we will use the set $\Theta_{3\text{-}RP}$ of the RPs of the prior distribution $\beta \sim N_2(\bar{\beta}, I_2)$. According to Theorem 1.2 in Fang and Wang [5], the set $\Theta_{3\text{-}RP}$ can be obtained by taking an inverse transformation of the following glp set in C^2,

$$\left\{ \left(\frac{1}{6}, \frac{3}{6}\right), \ \left(\frac{3}{6}, \frac{1}{6}\right), \ \left(\frac{5}{6}, \frac{5}{6}\right) \right\}.$$

The two sets $\Theta_{3\text{-}RP}$ and Θ_{HC} chosen from the prior distribution $N_2(\bar{\beta}, I_2)$ of β are shown in Table 9.1.

The optimal m-exact designs ($m = 8,16,24$) under the five optimality criteria in (9.9)–(9.13) for the first-order model (9.5) with random effects covariance matrix G_i ($i = 1, 2, 3$) are calculated numerically, where the sets $\Theta_{n\text{-}RP}$ and Θ_{HC} used in these criteria are given in Table 9.1. To save space, we only show the optimal 8-exact designs for the covariance matrix G_2 in Table 9.2.

It is observed from this table that for a given value of c, the designs have two support points except for the HCD-optimal designs on the cases $c = 0.01, 0.2$. The left endpoint of each design region is the common support of these designs, but the weights on it can be different.

In the following, taking for example, we make an efficiency comparison among the optimal 8-exact designs in the case of $c = 0.01$. We compute the D-efficiencies defined by (9.8) of the optimal 8-exact designs on the region $[0.01, 1]$ obtained under the criteria (9.9)–(9.13), respectively, with respect to each of the 100 locally D-optimal designs where the 100 values of β are randomly sampled from its prior distribution $N_2(\bar{\beta}, I_2)$. The results for each model with random effects covariance matrix G_j ($j = 1, 2, 3$) are shown in Figs. 9.1–9.3. In each plot, column 1 stands for the box plot of D-efficiency of the RPD-optimal design, column 2 for the box plot of

Table 9.1 The sets $\Theta_{3\text{-}RP}$ and Θ_{HC} for the prior distribution $\beta \sim N_2(\bar{\beta}, I_2)$

$\Theta_{3\text{-}RP}$	β_0	β_1	Θ_{HC}	β_0	β_1
β_{RP}^1	0.0326	−3.0000	β_{HC}^1	$1 - 1.96$	$-3 - 1.96$
β_{RP}^2	1.0000	−3.9674	β_{HC}^2	$1 - 1.96$	$-3 + 1.96$
β_{RP}^3	1.9674	−2.0326	β_{HC}^3	$1 + 1.96$	$-3 - 1.96$
			β_{HC}^4	$1 + 1.96$	$-3 + 1.96$

Table 9.2 The optimal 8-exact designs on $[c, 1]$ for the first-order model (9.5) with random effects covariance matrix $G_2 = 0.5\, I_2$ based on the sets $\Theta_{3\text{-}RP}$ and Θ_{HC} in Table 14.1

Criterion	$c = 0.01$	$c = 0.2$	$c = 0.4$
Local D	$\begin{pmatrix} 0.01 & 0.7250 \\ 0.3125 & 0.6875 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.8802 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 1 \\ 0.375 & 0.625 \end{pmatrix}$
HCD	$\begin{pmatrix} 0.01 & 0.5567 & 1 \\ 0.375 & 0.5 & 0.125 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.7611 & 1 \\ 0.375 & 0.5 & 0.125 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 0.9886 \\ 0.5 & 0.5 \end{pmatrix}$
HCMMD	$\begin{pmatrix} 0.01 & 0.4688 \\ 0.375 & 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.656 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 0.8524 \\ 0.5 & 0.5 \end{pmatrix}$
RPD	$\begin{pmatrix} 0.01 & 0.7075 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.9081 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 1 \\ 0.375 & 0.625 \end{pmatrix}$
RPMMD	$\begin{pmatrix} 0.01 & 0.8199 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 1 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 1 \\ 0.375 & 0.625 \end{pmatrix}$
Bayesian	$\begin{pmatrix} 0.01 & 0.7198 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.9185 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 1 \\ 0.5 & 0.5 \end{pmatrix}$

D-efficiency of the HCD-optimal design, column 3 for the box plot of D-efficiency of the Bayesian D-optimal design, column 4 for the box plot of D-efficiency of the RPMMD-optimal design, and column 5 stands for the box plot of D-efficiency of HCMMD-optimal design.

Figure 9.1 shows the results for the first-order Poisson model with random intercept. The median of the D-efficiency of the RPMMD-optimal design is the highest, even better than Bayesian optimal design, and the performance of the HCMMD-optimal design is the worst. Although the D-efficiency of the RPD-optimal design is a little lower than that of the HCD-optimal design, its median is above 0.8, which is acceptable in practice. Figures 9.2–9.3 show the results for the first-order Poisson model with both random intercept and random slope. These results show that the difference of the five designs shrinks, and their performances are comparable. It is noticed that the RPD-, RPMMD-and Bayesian D-optimal designs perform better than the HCD-and HCMMD-optimal designs. In conclusion, the optimality criteria based on the RPs is more efficient than that based on the hypercube method to overcome the problem of dependence of designs on the unknown parameters of the model.

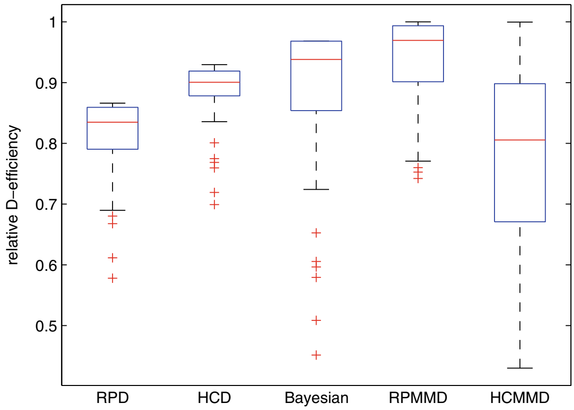

Fig. 9.1 Box plots of the D-efficiencies of the five optimal 8-exact designs with respect to the 100 locally D-optimal 8-exact designs on [0.01, 1] for the first-order model (9.5) with random effects covariance matrix G_1

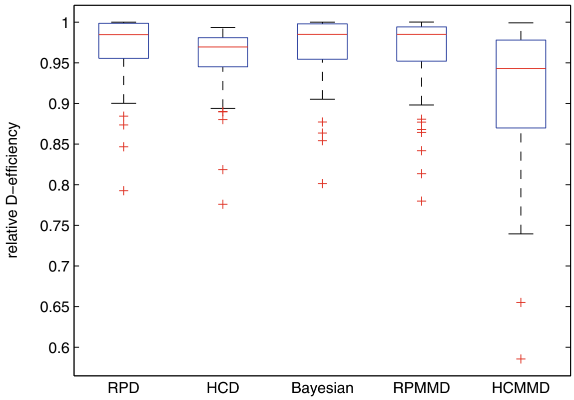

Fig. 9.2 Box plots of the D-efficiencies of the 8-exact optimal 8-exact designs with respect to the 100 locally D-optimal 8-exact designs on [0.01, 1] for the first-order model (9.5) with random effects covariance matrix G_2

Furthermore, we examine the affect of the number of RPs on the RPD- and RPMMD-optimal exact designs on [0.01, 1] for the first-order model (9.5) with random effects covariance matrix G_j ($j = 1, 2, 3$). The RPD-and RPMMD-optimal 8-exact designs are carried out under three sets $\Theta_{n\text{-}RP}$ ($n = 3, 5, 8$), and the D-efficiencies of these designs are calculated with respect to the locally D-optimal 8-exact designs at each of 100 values of β which are randomly sampled from the prior distribution $N_2(\bar{\beta}, I_2)$. For space reason, in Fig. 9.4 we only report a part of these D-efficiencies of the RPD-and RPMMD-optimal designs for the first-order model (9.5) with random effects covariance matrix G_1. These results show that the number of RPs has a slight impact on the RPD-and RPMMD-optimal designs.

9.4.1.2 The Case of Noncentral *t* Prior Distributions

We consider the case of non-normal prior distributions of the fixed effects. For illustration purpose, we assume that β_0 and β_1 in the first-order model (9.5) are indepen-

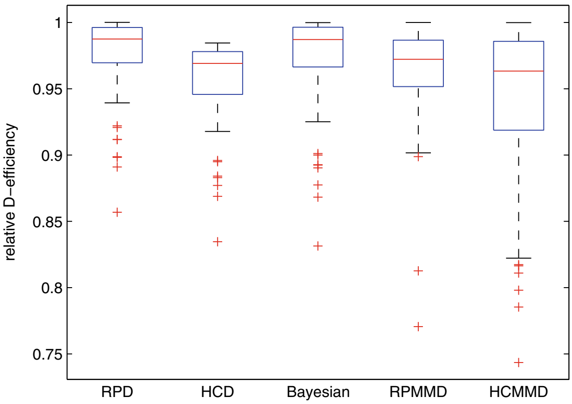

Fig. 9.3 Box plots of the D-efficiencies of the five optimal 8-exact designs with respect to the 100 locally D-optimal 8-exact designs on [0.01, 1] for the first-order model (9.5) with random effects covariance matrix G_3

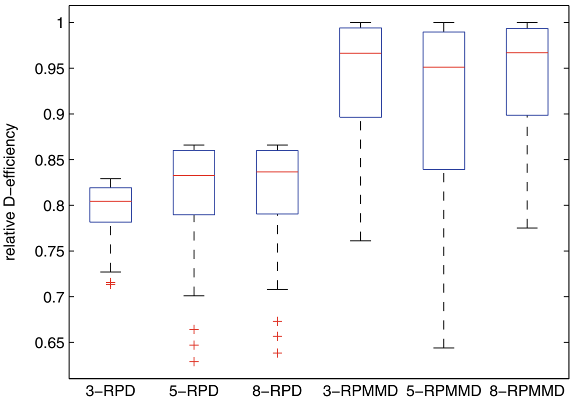

Fig. 9.4 Box plots of the D-efficiencies of the RPD-and RPMMD-optimal 8-exact designs under three sets $\Theta_{n\text{-}RP}$ ($n = 3, 5, 8$) with respect to the 100 locally D-optimal 8-exact designs on [0.01, 1] for the first-order model (9.5) with random effects covariance matrix G_1

dent and follow noncentral t distributions having means 1 and -3, respectively. Let $\beta_0 \sim t(q_0, \delta_0)$ and $\beta_1 \sim t(q_1, \delta_1)$. By assuming the degrees of freedom $q_0 = 4$ and $q_1 = 3$, the noncentrality parameters are then obtained by solving the equations

$$E(\beta_i) = \frac{\delta_i \Gamma(\frac{q_i - 1}{2})}{\Gamma(\frac{q_i}{2})} \sqrt{\frac{q_i}{2}}, \quad i = 0, 1,$$

which are $\delta_0 = 0.7979$ and $\delta_1 = -2.1708$, respectively. The set of RPs can also be obtained by the NTMs. Our computation is carried out in Matlab, and the sets $\Theta_{3\text{-}RP}$ and Θ_{HC} of β with this prior distribution are given in Table 9.3. The results in Table 9.4 are the optimal 8-exact designs under the five optimality criteria in (9.10)–(9.13) for the first-order model (9.5) with random effects covariance matrix $G_2 = 0.5\, I_2$ and the noncetral t prior distribution of β, where the sets $\Theta_{n\text{-}RP}$ and Θ_{HC} used in these criteria are as in Table 9.3. Compared with the results in Table 9.2, we observed that both the RPD-and RPMMD-optimal designs on the region $[c, 1]$ are very similar (except

Table 9.3 The sets $\Theta_{3\text{-}RP}$ and Θ_{HC} of $\beta = (\beta_0, \beta_1)^T$ whose components are independent and follow prior distributions $t(4, 0.7979)$ and $t(3, -2.1708)$ respectively

Θ_{RP}	β_0	β_1	Θ_{HC}	β_0	β_1
β_{RP}^1	−0.1823	−2.3957	β_{HC}^1	−1.4604	−9.4003
β_{RP}^2	0.8505	−4.4759	β_{HC}^2	−1.4604	−0.2209
β_{RP}^3	2.1199	−1.1984	β_{HC}^3	4.3557	−9.4003
			β_{HC}^4	4.3557	−0.2209

Table 9.4 The optimal 8-exact designs on $[c, 1]$ for the first-order model (9.5) with random effects covariance matrix $G_2 = 0.5\,I_2$ and the noncentral t prior distribution of β, based on the sets $\Theta_{3\text{-}RP}$ and Θ_{HC} in Table 9.3

criterion	$c = 0.01$	$c = 0.2$	$c = 0.4$
HCD	$\begin{pmatrix} 0.01 & 0.2456 & 1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.4402 & 1 \\ 0.375 & 0.5 & 0.125 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 0.6527 & 1 \\ 0.375 & 0.5 & 0.125 \end{pmatrix}$
HCMMD	$\begin{pmatrix} 0.01 & 0.24 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.4209 \\ 0.5 & 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 0.6208 \\ 0.5 & 0.5 \end{pmatrix}$
RPD	$\begin{pmatrix} 0.01 & 0.6979 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.8862 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 1 \\ 0.375 & 0.625 \end{pmatrix}$
RPMMD	$\begin{pmatrix} 0.01 & 0.7204 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.7420 \\ 0.375 & 0.625 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 0.9498 \\ 0.375 & 0.625 \end{pmatrix}$
Bayesian	$\begin{pmatrix} 0.01 & 0.7110 & 1 \\ 0.375 & 0.25 & 0.125 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.8658 & 1 \\ 0.5 & 0.25 & 0.25 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 0.9648 & 1 \\ 0.5 & 0.125 & 0.375 \end{pmatrix}$

RPMMD at $c = 0.2$), while others are much different, based on the two kinds of the prior distribution of β.

9.4.2 Designs for the Second-Order Poisson Mixed Model

A similar discussion to the previous subsection is done to the second-order Poisson mixed model in (9.6). For illustration, we assume that the covariance matrix of the random effects $b = (b_0, b_1, b_2)^T$ is $G = 0.5\,I_3$, and the prior distribution of the fixed effects $\beta = (\beta_0, \beta_1, \beta_2)^T$ is normal distribution with mean $\bar{\beta} = (1, -3, -0.9)^T$ and covariance matrix I_3. In this case, the percentile set Θ_{HC} contains 8 points, and for comparison we choose the set $\Theta_{7\text{-}RP}$ having seven RPs of the prior distribution $\beta \sim N_3(\bar{\beta}, I_3)$, which is obtained by the inverse transformation method from the following glp set,

$$\Big\{\Big(\frac{1}{14}, \frac{5}{14}, \frac{9}{14}\Big), \Big(\frac{3}{14}, \frac{11}{14}, \frac{5}{14}\Big), \Big(\frac{5}{14}, \frac{3}{14}, \frac{1}{14}\Big), \Big(\frac{7}{14}, \frac{9}{14}, \frac{11}{14}\Big), \Big(\frac{9}{14}, \frac{1}{14}, \frac{7}{14}\Big), \Big(\frac{11}{14}, \frac{7}{14}, \frac{3}{14}\Big), \Big(\frac{13}{14}, \frac{13}{14}, \frac{13}{14}\Big)\Big\}.$$

The two sets $\Theta_{7\text{-}RP}$ and Θ_{HC} are shown in Table 9.5.

Table 9.6 shows the six kinds of optimal 8-exact designs on the region $[c, 1]$ with $c = 0.01, 0.2$ for the second-order model (9.6) with the random effects covariance matrix $Cov(b) = 0.5 I_3$. These designs are obtained numerically under the six optimality criteria given in (9.7), (9.9)–(9.13), where the sets $\Theta_{7\text{-}RP}$ and Θ_{HC} in Table 9.5 are used in (9.10)–(9.13) correspondingly.

As in the previous subsection, we are going to make a comparison among these designs. We generate randomly 100 values of β from the prior distribution $\beta \sim N_3(\bar{\beta}, I_3)$, and find out the locally D-optimal 8-exact designs on the region $[0.01, 1]$ at each of these values of β. Then we calculate the D-efficiencies of the RPD-, HCD-, Bayesian D-, RPMMD-and HCMMD-optimal 8-exact designs relative to each of these locally D-optimal designs. The box plots of these D-efficiencies are shown in Fig. 9.5.

As shown in Fig. 9.5, the medians of D-efficiencies of the RPD-, HCD-, Bayesian D-, RPMMD-optimal designs are all greater than 0.95, while the median of D-efficiencies of the HCMMD-optimal design is 0.8. The performance of the RPD-optimal design is slightly better than the HCD-and Bayesian D-optimal designs. And the performance of the RPMMD-optimal design is much better than the HCMMD-optimal design.

Table 9.5 The sets $\Theta_{7\text{-}RP}$ and Θ_{HC} for the prior distribution $\beta \sim N_3(\bar{\beta}, I_3)$

$\Theta_{7\text{-}RP}$	β_0	β_1	β_2	Θ_{HC}	β_0	β_1	β_2
β_{RP}^1	−0.4652	−3.3661	−0.5339	β_{HC}^1	$1 - 1.96$	$-3 - 1.96$	$-0.9 - 1.96$
β_{RP}^2	0.2084	−2.2084	−1.2661	β_{HC}^2	$1 - 1.96$	$-3 + 1.96$	$-0.9 - 1.96$
β_{RP}^3	0.6339	−3.7916	−2.3652	β_{HC}^3	$1 - 1.96$	$-3 - 1.96$	$-0.9 + 1.96$
β_{RP}^4	1.0000	−2.6339	−0.1084	β_{HC}^4	$1 - 1.96$	$-3 + 1.96$	$-0.9 + 1.96$
β_{RP}^5	1.3661	−4.4652	−0.9	β_{HC}^5	$1 + 1.96$	$-3 - 1.96$	$-0.9 - 1.96$
β_{RP}^6	1.7916	−3	−1.6916	β_{HC}^6	$1 + 1.96$	$-3 + 1.96$	$-0.9 - 1.96$
β_{RP}^7	2.4652	−1.5348	0.5652	β_{HC}^7	$1 + 1.96$	$-3 - 1.96$	$-0.9 + 1.96$
				β_{HC}^8	$1 + 1.96$	$-3 + 1.96$	$-0.9 + 1.96$

Table 9.6 The optimal 8-exact designs on $[c, 1]$ for the second-order model (9.6) with the random effects covariance matrix $Cov(b) = 0.5\, I_3$

criterion	$c = 0.01$	$c = 0.2$
Local D	$\begin{pmatrix} 0.01 & 0.3365 & 0.9665 \\ 0.25 & 0.375 & 0.375 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.4923 & 1 \\ 0.25 & 0.375 & 0.375 \end{pmatrix}$
HCD	$\begin{pmatrix} 0.01 & 0.2807 & 0.7627 & 1 \\ 0.25 & 0.375 & 0.25 & 0.125 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.4582 & 0.8337 & 1 \\ 0.25 & 0.375 & 0.125 & 0.25 \end{pmatrix}$
RPD	$\begin{pmatrix} 0.01 & 0.3186 & 0.8872 & 1 \\ 0.25 & 0.375 & 0.25 & 0.125 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.4772 & 0.9316 & 1 \\ 0.25 & 0.375 & 0.125 & 0.25 \end{pmatrix}$
HCMMD	$\begin{pmatrix} 0.01 & 0.2122 & 0.6354 \\ 0.25 & 0.375 & 0.375 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.3868 & 0.8163 \\ 0.25 & 0.375 & 0.375 \end{pmatrix}$
RPMMD	$\begin{pmatrix} 0.01 & 0.3412 & 1 \\ 0.25 & 0.375 & 0.375 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.4644 & 1 \\ 0.25 & 0.375 & 0.375 \end{pmatrix}$
Bayesian D	$\begin{pmatrix} 0.01 & 0.3219 & 0.8124 & 0.9686 \\ 0.25 & 0.375 & 0.25 & 0.125 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0.4766 & 0.9905 \\ 0.5 & 0.375 & 0.125 \end{pmatrix}$

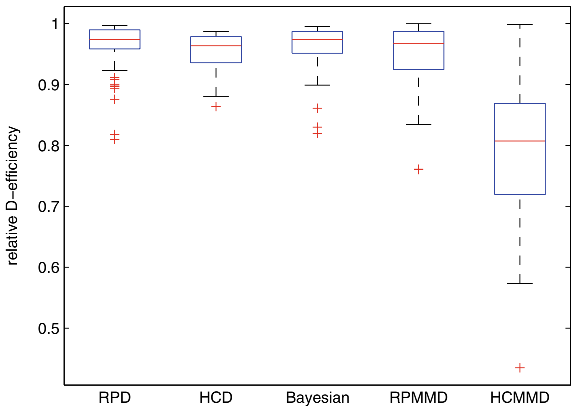

Fig. 9.5 Box plots of the D-efficiencies of the five optimal 8-exact designs relative to the 100 locally D-optimal 8-exact designs on $[0.01, 1]$ for the second-order model (9.6) with random effects covariance matrix $Cov(b) = 0.5\, I_3$

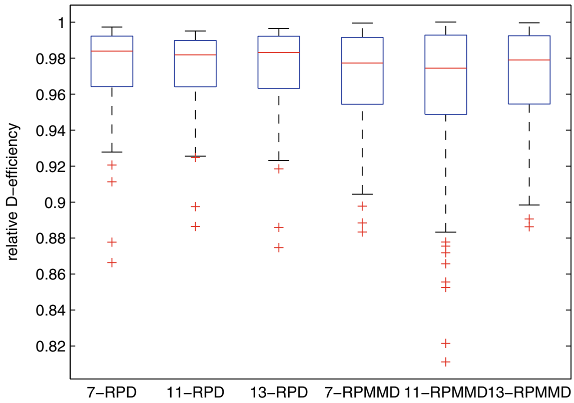

Fig. 9.6 Box plots of the D-efficiencies of the RPD-and RPMMD-optimal 8-exact designs under three sets $\Theta_{7\text{-}RP}$,$\Theta_{11\text{-}RP}$, $\Theta_{13\text{-}RP}$ with respect to the 100 locally D-optimal 8-exact designs on [0.01, 1] for the second-order model (9.6) with random effects covariance matrix $Cov(b) = 0.5\, I_3$

We here examine the affect of the number of RPs used in the RPD- and RPMMD-optimality criteria on the optimal designs for the model (9.6). The RPD-and RPMMD-optimal 8-exact designs on $[0.01, 1]$ under three sets $\Theta_{n\text{-}RP}$ $(n = 7, 11, 13)$ are calculated numerically, and the D-efficiencies of these designs are computed with respect to the locally D-optimal 8-exact designs at each of the 100 values of β which are randomly sampled from the prior distribution $N_3(\bar{\beta}, I_3)$. These results in Fig. 9.6 show that the number of RPs has a slight impact on the RPD-and RPMMD-optimal designs.

9.5 Concluding Remarks

This paper concerns with optimal and robust design problems for Poisson mixed models. Two optimality criteria, termed RPD-optimality and RPMMD-optimality, for the Poisson mixed model are introduced by using the RPs of the prior distribution of fixed effects. The purpose of these two criteria is to overcome the dependence problem of D-optimality on the values of unknown parameters. By assuming the prior distribution of fixed effects is a multivaraite normal distribution with independent components, we obtain the RPs by using the transformation method. The numerical results for the first-and second-order models show that the optimal designs based on the RPs are more robust than those based on the hypercube method. Moreover, the number of RPs has a slight impact on both RPD-and RPMMD-optimal designs. Therefore, a small number of RPs used in the RPD-and RPMMD-optimality criteria may yield a good robustness against parameter uncertainty. Hence, our results will give more options to the experimenters.

In aspects of computation, the running times of constructing the RPD-and RPMMD-optimal designs are much less than that of the HCD-and HCMMD-optimal designs, respectively. The computation time of constructing the Bayesian D-optimal design is much longer than others due to the long time required in computation of the Bayesian criterion for a given design.

Moreover, in our computation the prior distribution of the fixed effects is assumed to have independent components, and then the RPs are obtained by using the inverse transformation method. If the prior distributions of the fixed effects have correlated components, the RPs can be generated by other methods proposed in, e.g., [5, 18, 19] and the references therein.

Acknowledgments The work is supported by the NSFC grants 11971318, 11871143.

References

1. Atkinson, A.C., Cox, D.C.: Planning experiments for discriminating between models. J. Roy. Statist. Soc. Ser. B **36**, 321–348 (1974)
2. Atkinson, A.C., Donev, A.N., Tobias, R.D.: Optimal Experimental Designs, With SAS. Oxford university Press, New York (2007)
3. Bohachevsky, I.O., Johnson, M.E., Stein, M.L.: Generalized simulated annealing for function optimization. Technometrics **28**, 209–217 (1986)
4. Dette, H., Haines, L.M., Imhof, L.A.: Maximin and Bayesian optimal designs for regression models. Statist. Sinica **17**, 463–480 (2007)
5. Fang, K.T., Wang, Y.: Number-Theoretic Methods in Statistics. Chapman and Hall, London (1994)
6. Haines, L.M.: The application of the annealing algorithm to the construction of exact optimal designs for linear regression models. Technometrics **29**, 439–447 (1987)
7. Haines, L.M.: A geometric approach to optimal design for one-parameter non-linear models. J. Roy. Statist. Soc. Ser. B **57**, 575–598 (1995)
8. Khuri, A.I., Mukherjee, B., Sinha, B.K., Ghosh, M.: Design issue for generalized linear models: a review. Statist. Sci. **21**, 376–399 (2006)
9. Foo, L.K., Duffull, S.: Methods of robust design of nonlinear models with an application to pharmacokinetics. J. Biopharm. Statist. **20**, 886–902 (2010)
10. McCullagh, P.M., Nelder, J.A.S.: Generalized Linear Models, 2nd edn. Chapman and Hall, London (1989)
11. McGree, J.M., Duffull, S.B., Eccleston, J.A.: Optimal design for studying bioimpedance. Physiol. Meas. **28**, 1465–1483 (2007)
12. Niaparast, M.: On optimal design for a Poisson regression model with random intercept. Statist. Probab. Lett. **79**, 741–747 (2009)
13. Niaparast, M.: On optimal design for mixed effects Poisson regression models. Ph.D. thesis, Otto-von-Guericke University, Magdeburg (2010)
14. Niaparast, M., Schwabe, R.: Optimal design for quasi-likelihood estimation in Poisson regression with random coefficients. J. Statist. Plann. Inference **143**, 296–306 (2013)
15. Wang, Y.P.: Optimal experimental designs for the Poisson regression model in toxicity studies. Ph.D. thesis, Virginia Polytechnic Institute and State University (2002)
16. Wedderburn, R.W.M.: Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. Biometrika **61**, 439–447 (1974)
17. Wong, W.K.: A unified approach to the construction of minimax designs. Biometrika **79**, 611–619 (1992)

18. Zhou, M., Wang, W.J.: Representative points of Student's t_n distribution and their application in statistical simulation (in Chinese). Acta. Math. Appl. Sinica. Chin. Ser. **39**, 620–640 (2016)
19. Zhou, Y.D., Fang, K.T.: FM-criterion for representative points (in Chinese). Sci. Chin. Math. **49**, 1009–1020 (2019)