Journal of Statistical Modelling: Theory and Applications
Vol. 4, No. 1, 2023, pp. 147-156
Yazd University Press 2023

JSMTA

Research Paper

E-optimal designs for Poisson regression models with random coefficients

Mehrdad Niaparast*, Sahar Mehr Mansour, Akram Zangenehmehr
Department of Statistics, Razi University, Kermanshah, Iran

Received: February 14, 2024/ Revised: April 12, 2024/ Accepted: April 17, 2024

Abstract: Most of the research on optimal designs concentrates on D-optimal designs for linear and nonlinear models with fixed effects. Recently nonlinear models with random effects have been of great attention because these models are more applicable to describing real data. In this paper, E-optimal designs for the Poisson regression with random effects have been considered. A new version of the equivalence theorem is prepared for this criterion in the Poisson regression model with random effects.

Keywords: E-optimal design; Poisson regression model; Quasi-likelihood method; Random effect.
Mathematics Subject Classification (2010): 92K05.

1 Introduction

Experimental design is a powerful statistical tool for scientific research and industrial applications. A well-planned experimentation is an effective way to improve the quality of the analysis. The idea of optimal designs is to find the best values of the control variables, which maximize the amount of information obtained from observations. To compare different designs, it is necessary to define a criterion, a real-valued function on the Fisher information matrix for the parameter vector that equals the inverse of the variance-covariance matrix. Most of the literature on the optimal design concentrates strongly on D-optimality, which is the based on minimizing of the real value function, determinant , of the variance-covariance matrix of estimators. In contrast, it has paid less attention to E-optimal designs, which minimize the largest eigenvalue of the variance-covariance matrix of estimators. E-criterion is recognized as special case of the Kiefer's Φ_q-criterion for $q \to \infty$, for example see Kiefer (1974).

*Corresponding author: Niaparast@razi.ac.ir

Dette and Studen (1993) consider the E-optimal design problem by investigating the geometric properties of linear models. They presented new proof for spring balance and chemical balance weighting designs. For the linear and nonlinear models with two parameters, the optimality of their design was investigated analytically by Dette and Haines (1994). The E-optimal designs for nonlinear regression models are presented by Dette and Melas (2004) in detail. They showed the optimality that occurred in the Chebyshev points. Prus (2019) discussed the equivalence of the E-criterion between the fixed effects model and random coefficients regression models.

This work concentrates on the E-optimal designs for the Poisson regression model with random effects. This model is widely used in biosciences when replicated measurements are available from different individuals, for example Verbeke and Molenberghs (2000). Due to the random effects, an explicit form for the variance-covariance matrix could not be achieved. Niaparast (2009) and Niaparast and Schwabe (2013) used the quasi-likelihood method and obtained a quasi-information matrix. Then consider D-optimal designs for the quasi-likelihood estimator of parameters in mixed effects Poisson regression models. We use these results to obtain candidates for the E-optimal designs and check the optimality of these designs by using equivalence theorems. Furthermore, among different strategies to find the optimal design for this model, we consider designs as local because it is the basis for others.

This paper is organized as follows. In Section 2, we will introduce the model and the variance-covariance structure of the quasi-likelihood estimator of parameters in the Poisson regression model with random effects. In our case, a new equivalence theorem for the E-criterion is deferred in Section 3. We will also obtain E-optimal designs for some special cases of the Poisson regression model with random effects.

2 A review on the structure of model and design specification

In this paper, we consider a Poisson regression model with random coefficients in which the j-th observation of individual i, Y_{ij}, is distributed as a Poisson distribution with the natural log link, $\log(\lambda_{ij}) = \boldsymbol{f}^T(x_{ij})\boldsymbol{b}_i$, for $i = 1, \cdots, n$ and $j = 1, \cdots, m_i$. Here n is the number of individuals and m_i is the number of observations per individual, $\boldsymbol{f} = (f_0, f_1, \cdots, f_{p-1})^T$ is the known regression function and the $p \times 1$ vector $\boldsymbol{b}_i = (b_{0,i}, \cdots, b_{p-1,i})^T$ is a vector of random effects which is normally distributed with the mean vector $\boldsymbol{\beta} = (\beta_0, \cdots, \beta_{p-1})^T$ and known variance-covariance matrix $\boldsymbol{\Sigma}$. Further, the random effects vector and observations in different individuals are assumed to be uncorrelated, whereas observations into individuals are correlated.

On the individual level, the vector of the all m_i observations, $\boldsymbol{Y}_i = (Y_{i1}, \cdots, Y_{im_i})^T$, has the mean $E(\boldsymbol{Y}_i) = (\mu(x_{ij}), \cdots, \mu(x_{im_i}))^T$ where $\mu(x) = \exp(\boldsymbol{f}^T(x)\boldsymbol{\beta} + \frac{1}{2}\sigma(x; x'))$ is the mean function with the dispersion function $\sigma(x; x') = \boldsymbol{f}^T(x)\boldsymbol{\Sigma}\boldsymbol{f}(x')$. Also the variance-covariance matrix structure of $\boldsymbol{Y}_i$ is given by $\mathrm{Var}(\boldsymbol{Y}_i) = \boldsymbol{A}_i + \boldsymbol{A}_i\boldsymbol{C}_i\boldsymbol{A}_i$ where $\boldsymbol{A}_i = \mathrm{diag}\{\mu(x_{ij})\}_{j=1,\cdots,m_i}$ is a diagonal matrix with entries $\mu(x_{ij})$ for $j = 1, \cdots, m_i$ and $\boldsymbol{C}_i = (c(x_{ij}; x_{ik}))_{j,k=1,\cdots,m_i}$. Here $c(x; x') = \exp(\sigma(x; x')) - 1$ is the variance correction term.

By the same way, on the population level, the vector of the all observations $\boldsymbol{Y} =$

$(\boldsymbol{Y}_1^T, \cdots, \boldsymbol{Y}_n^T)^T$ has the mean $E(\boldsymbol{Y}) = (E^T(\boldsymbol{Y}_1), \cdots, E^T(\boldsymbol{Y}_n))^T$ and the variance-covariance matrix structure of $\mathrm{Var}(\boldsymbol{Y}) = \boldsymbol{A} + \boldsymbol{ACA}$ where $\boldsymbol{A} = \mathrm{diag}\{\boldsymbol{A}_i\}_{i=1,\cdots,n}$ and $\boldsymbol{C} = \mathrm{diag}\{\boldsymbol{C}_i\}_{i=1,\cdots,n}$.
Niaparast and Schwabe (2013) obtained the quasi-information matrix for the vector of fixed effect parameters, $\boldsymbol{\beta}$, in the Poisson regression model with random effects as follows,

$$\mathfrak{M}(\boldsymbol{\beta}) = \sum_i \mathfrak{M}_i(\boldsymbol{\beta}) = \sum_i \boldsymbol{F}_i^T(\boldsymbol{A}_i^{-1} + \boldsymbol{C}_i)^{-1}\boldsymbol{F}_i,$$

where $\boldsymbol{F}_i = (\boldsymbol{f}(x_{i1}), \cdots, \boldsymbol{f}(x_{im_i}))^T$.
Let $\xi_i = \left\{ \begin{array}{c} x_{i1}, \cdots, x_{is_i} \\ p_{i1}, \cdots, p_{is_i} \end{array} \right\}$ be a design that we have planned to conduct experiments for ith individual where $p_j = \frac{m_{ij}}{m_i}$ $(j = 1, \cdots, s_i)$ stands for the proportion of observations taken at x_j and $\sum_j m_{ij} = m_i$. Also, denote the population design by $\zeta = \left\{ \begin{array}{c} \xi_1, \cdots, \xi_r \\ w_1, \cdots, w_r \end{array} \right\}$ where $w_i = \frac{n_i}{n}$ $(i = 1, \cdots, r)$ is the proportion of individuals observed under the same design ξ_i where $\sum n_i = n$. Niaparast and Schwabe (2013) have discussed the optimality of designs issue to obtain an optimal design ζ^*. They showed to discover ζ^* it is sufficient to consider single-group design instead of population design, in other words $\zeta^* = \left\{ \begin{array}{c} \xi^* \\ 1 \end{array} \right\}$. Therefore, we can observe all individuals under the same design ξ. For design ξ, the quasi-information matrix can exhibit as follows,

$$\mathfrak{M}_{\boldsymbol{\beta}}(\xi) = \boldsymbol{F}_\xi^T(\boldsymbol{A}_\xi^{-1} + \boldsymbol{C}_\xi)^{-1}\boldsymbol{F}_\xi,$$

where

$$\boldsymbol{F}_\xi = (\boldsymbol{f}(x_1), \cdots, \boldsymbol{f}(x_s))^T, \quad \boldsymbol{A}_\xi = \mathrm{diag}\{m_j\mu(x_j)\}_{j=1,\cdots,s}, \quad \boldsymbol{C}_\xi = (c(x_j, x_k))_{j,k=1,\cdots,s}.$$

Note that the quasi-information matrix depends on the unknown parameter vector $\boldsymbol{\beta}$ through the mean function. Such a dependency usually occurs in nonlinear models. To cover this difficulty, we will apply a locally optimal design approach that uses only a point estimation.

3 Locally E-optimal

In optimal experimental design, it is necessary to determine the criterion which measures the quality of a design. Among different criteria in the literature for this purpose, we concentrate on the E-optimal criterion, which focuses on the eigenvalue of the information matrix, for example see Pukelsheim (2006). We define the E-criterion

$$\Psi(\xi) = \lambda_{max}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi)),$$

as the largest eigenvalue of the inverse of the quasi informatin matrix. Hense, ξ^* is E-optimal if that minimizes $\Psi(\xi)$. For the E-criterion, monotonicity can be seen easily concerning loewner ordering. The following lemma verifies concavity for the E-criterion.

Lemma 3.1. *The E-criterion is a concave function of ξ on the set of all designs.*

Proof. The concavity of the E-criterion follows directly from Lemma 5.1 in Niaparast and Schwabe (2013). Then the result is obtained in the sense of the Loewner ordering of nonnegative definiteness. The following theorem gives a sufficient and necessary condition that ξ^* is E-optimal. For stating the equivalence theorem, we need to introduce the sensitivity function $d(x,\xi)$ as

$$\begin{aligned} d(x,\xi) &= m\mu(x)\ (\boldsymbol{f}(x) - \boldsymbol{F}_\xi^T(\boldsymbol{A}_\xi^{-1} + \boldsymbol{C}_\xi)^{-1}\boldsymbol{c}_{\xi,x})^T\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi)\boldsymbol{q}\boldsymbol{q}^T \\ &\quad \times\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi)(\boldsymbol{f}(x) - \boldsymbol{F}_\xi^T(\boldsymbol{A}_\xi^{-1} + \boldsymbol{C}_\xi)^{-1}\boldsymbol{c}_{\xi,x}), \end{aligned}$$

where $\boldsymbol{q}$ is the eigenvector corresponding to $\lambda_{max}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))$ and the vector $c_{\xi,x} = (c(x_j,x))_{j=1,\dots,s}$ of joint (correlation) correction terms for the settings $x_1,\dots,x_s$ of a design ξ for prediction of a further setting x. □

Theorem 3.2. *Let ξ^* is a design with the maximum eigenvalue $\lambda_{max}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))$ of the inverse of quasi-information matrix with multiplicity 1. ξ^* is locally E-optimal at $\boldsymbol{\beta}$ in the mixed effects Poisson regression model, if and only if for all $x \in \mathcal{X}$*

$$\begin{aligned} d(x,\xi^*) &\le \boldsymbol{q}^T(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*))\boldsymbol{q} - \boldsymbol{q}^T(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*)\boldsymbol{F}_{\xi^*}^T(\boldsymbol{A}_{\xi^*}^{-1} + \boldsymbol{C}_{\xi^*})^{-1}\boldsymbol{C}_{\xi^*} \\ &\quad \times(\boldsymbol{A}_{\xi^*}^{-1} + \boldsymbol{C}_{\xi^*})^{-1}\boldsymbol{F}_{\xi^*}\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*))\boldsymbol{q}. \end{aligned}$$

The sensitivity function $d(x,\xi^*)$ of the optimal design ξ^* attains its maximum at the support of ξ^*. Moreover, equality hold for all support points of ξ^*.

The proof can be found in Appendix.

4 Application

We are going to investigate the simple Poisson regression model in two cases, random intercept and random slope, respectively:

$$\begin{aligned} \lambda &= \exp(b_0 + \beta_1 x) \text{ with } b_0 \sim N(\beta_0, \sigma^2), \\ \lambda &= \exp(\beta_0 + b_1 x) \text{ with } b_1 \sim N(\beta_1, \sigma^2), \end{aligned}$$

where $\boldsymbol{\beta} = (\beta_0,\beta_1)^T$ and $\boldsymbol{f(x)} = (1,x)^T$. In this case $\mu(x) = \exp(\beta_0 + \beta_1 x + \frac{1}{2}\sigma(x;x'))$ with $\sigma(x;x') = (1,x)\boldsymbol{\Sigma}(1,x)^T$. Note that we restrict our research to designs with only two different settings x_1 and x_2.

Lemma 4.1. *If ξ is any design with two points in the experimental setting, i.e. $\xi = \left\{ \begin{array}{cc} x_1 & x_2 \\ p & (1-p) \end{array} \right\}$, then the largest eigenvalue of inverse quasi-information matrix for β is as*

$$\lambda_{max}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi)) = \frac{\mathrm{tr}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi)) + \sqrt{(\mathrm{tr}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi)))^2 - 4\det(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))}}{2}.$$

Proof. By standard definition of eigenvalues, we can write a quadratic polynomial as

$$\lambda^2 - \lambda(\mathrm{tr}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))) + \det(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi)) = 0.$$

The result is obtaind by solving this equation. □

Then we obtain the following result by the above criterion definition to check that the two point design is optimal.

Corollary 4.2. *Design ξ^* is the E-optimal design for two points design in the Poisson regression model with random coefficient if and only if for all $\boldsymbol{x} \in \mathcal{X}$*

$$\begin{aligned} d(x,\xi^*) \quad \leq \quad & \mathrm{tr}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*)) - \mathrm{tr}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*)D\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*)) + \frac{1}{B}\big[\mathrm{tr}^2(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*) \\ & -\mathrm{tr}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*))\mathrm{tr}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*)D\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*))\big] - \frac{2\det(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*))}{B} \\ & \times(2-\mathrm{tr}(D\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*))), \end{aligned}$$

where $B = \sqrt{\mathrm{tr}^2(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*)) - 4\det(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*))}$ *and* $D = \boldsymbol{F}_{\xi^*}^T(\boldsymbol{A}_{\xi^*}^{-1} + \boldsymbol{C}_{\xi^*})^{-1}\boldsymbol{C}_{\xi^*}(\boldsymbol{A}_{\xi^*}^{-1} + \boldsymbol{C}_{\xi^*})^{-1}\boldsymbol{F}_{\xi^*}$ *and also the sensitivity function is*

$$\begin{aligned} d(x,\xi^*) \quad = \quad & m\mu(x)R^T\mathfrak{M}_{\boldsymbol{\beta}}^{-2}(\xi^*)R \\ & +\frac{1}{B}[m\mu(x)\mathrm{tr}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*))R^T\mathfrak{M}_{\boldsymbol{\beta}}^{-2}(\xi^*)R - 2m\mu(x)\det(\mathfrak{M}_{\boldsymbol{\beta}}^{-1})R^T\mathfrak{M}_{\boldsymbol{\beta}}^{-1}R, \end{aligned}$$

with $R = \boldsymbol{f}(x) - \boldsymbol{F}_{\xi^*}^T(\boldsymbol{A}_{\xi^*}^{-1} + \boldsymbol{C}_{\xi^*})^{-1}\boldsymbol{c}_{\xi^*,x}$.

In most situations, the explanatory variables explain nonnegative quantities. In particular, an experimenter may define the design region as $\mathfrak{X} = [h,g]$, where both h and g are nonnegative. Furthermore, we consider the case that μ_j is a monotone function of x_j especially where μ_j is a decreasing function of x_j. Hence, it is reasonable to define the canonical standardized mean $\tilde{\mu_j} = \tilde{\mu}(x_j) = \frac{\mu(x_j)}{\mu(h)}$ will always lie in $[\tilde{\mu_g}, 1]$ corresponding to $[h,g]$. In this paper, we set $[h,g] = [0,\infty)$.

4.1 The case random intercept

It was shown by Niaparast (2009), Lemma 4.1.3, that $\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi)$ for a Poisson regression model with random intercept is

$$\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi) = \boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi) + \boldsymbol{U},$$

where $\boldsymbol{M}_{\boldsymbol{\beta}}(\xi) = \exp(\frac{1}{2}\sigma^2)\boldsymbol{F}_{\xi}^T\tilde{\boldsymbol{A}}_{\xi}\boldsymbol{F}_{\xi}$ with $\tilde{\boldsymbol{A}}_{\xi} = \mathrm{diag}\{m_j\exp(\boldsymbol{f}^T(x_j)\boldsymbol{\beta})\}_{j=1,2}$. Noth that $\boldsymbol{F}_{\xi}^T\tilde{\boldsymbol{A}}_{\xi}\boldsymbol{F}_{\xi}$ is information matrix for the corresponding model without random effects and $\boldsymbol{U} = (\exp(\sigma^2)-1)\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Calculations show that for $\boldsymbol{\beta} = (-2,-5)^T$ and $m = 100$, we have $\tilde{\mu}_1^* = 0.0775$ and $\tilde{\mu}_2^* = 1$, whereas optimal weights are different in terms of σ^2. Since $\tilde{\mu}^* = \exp(\beta_1 x^*)$, we can drive optimal support points with $x_i^* = \frac{\log(\tilde{\mu}_i^*)}{\beta_1}$ for $i = 1,2$.

Figure 1 indicates optimal weights for some special values of σ^2. The results show as σ^2 increases, the optimal weight decreases. It is easy to see that the equivalence theorem can be indicated as following corollary.

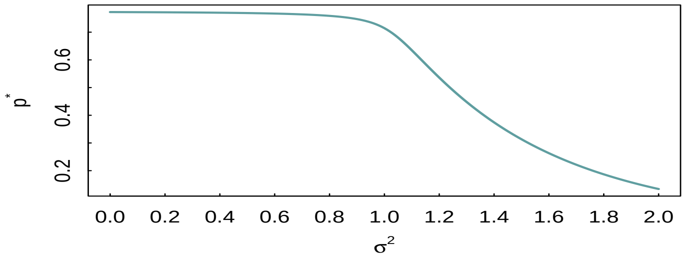

Figure 1: Optimal weights p^* for model with random intercept for different values of σ^2.

Corollary 4.3. *The design ξ^* is E-optimal design for a Poisson regression model with random intercept if and only if for all $x \in \mathcal{X}$*

$$\begin{aligned} d(x,\xi^*) \quad \leq \quad & \mathrm{tr}(\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*)) + \frac{1}{B}\big[\mathrm{tr}(\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*))\mathrm{tr}(\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*) + \boldsymbol{U}) \\ & -2\det(\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*) + \boldsymbol{U})\mathrm{tr}(\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*) + \boldsymbol{U})^{-1}\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*)), \end{aligned}$$

where $B = \sqrt{\mathrm{tr}^2(\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*) + \boldsymbol{U}) - 4\det(\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*) + \boldsymbol{U})}$ *and*

$$\begin{aligned} d(x,\xi^*) \quad = \quad & m\mu(x)\boldsymbol{f}^T(x)\boldsymbol{M}_{\boldsymbol{\beta}}^{-2}(\xi^*)\boldsymbol{f}(x) + \frac{1}{B}\big[\mathrm{tr}(\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*) + \boldsymbol{U}) \\ & \times(m\mu(x)\boldsymbol{f}^T(x)\boldsymbol{M}_{\boldsymbol{\beta}}^{-2}(\xi^*)\boldsymbol{f}(x)) - 2\det(\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*) + \boldsymbol{U}) \\ & \times(m\mu(x)\boldsymbol{f}^T(x)\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*)(\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*) + \boldsymbol{U})^{-1}\boldsymbol{M}_{\boldsymbol{\beta}}^{-1}(\xi^*)\boldsymbol{f}(x))\big], \end{aligned}$$

is the sensitivity function.

Figure 2 confirm the optimality ξ^* by using above colollary.

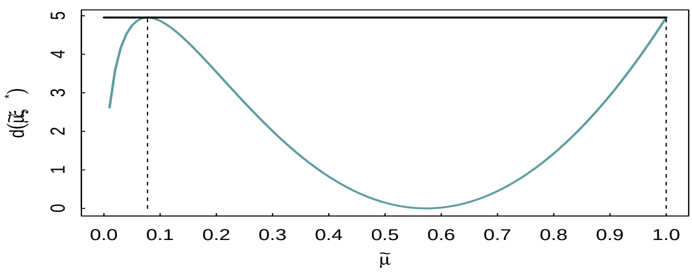

Figure 2: Sensitivity function of ξ^* with $\tilde{\mu}_1^* = 0.0775$ and $\tilde{\mu}_2^* = 1$.

As a benchmark, we turn to consider the efficiency of a given design with respect to the locally E-optimal design using E_{eff} which is defined by

$$\mathrm{E}_{eff} = \frac{\lambda_{max}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*))}{\lambda_{max}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))}.$$

Usually in practice, a standard two points design with uniform weights and endpoints $\tilde{\mu} = 0.001$ which is near zero, and $\tilde{\mu} = 1$ is used to carry out experiments. In Table

1 we provide comparsion of such design with E-optimal designs. It is shown that the efficiency of standard design tend to be closer to E-optimal design when the value of variance increase. This trend meight be correct when the parameters β_0 and β_1 are increase.

Table 1: Efficiency for model with random intercept.

σ^2	$\mathrm{E}_{eff}(\beta_0=-3,\beta_1=-5)$	$\mathrm{E}_{eff}(\beta_0=-2,\beta_1=-5)$	$\mathrm{E}_{eff}(\beta_0=-2,\beta_1=-2)$
0	0.076	0.077	0.083
0.5	0.078	0.078	0.146
1	0.078	0.084	0.464
1.5	0.083	0.192	0.995

4.2 The case random slope

Unlike the random intercept model, the random effect's size depends on the explanatory variable's value in the random slope model. For this model, the corresponding mean function equals $\mu = \exp(\beta_0 + \beta_1 x + \frac{1}{2}\sigma^2 x^2)$. So, the canonical standardized mean $\tilde{\mu} = \exp(\beta_1 x + \frac{1}{2}\sigma^2 x^2)$. Recall that we have supposed that the mean response is a decreasing function of the explanatory variable. It means that

$$\frac{d\mu}{dx} = (\beta_1 + \sigma^2 x)\mu < 0,$$

and then $\beta_1 + \sigma^2 x < 0$. This leads to a restricted region for points of design, i.e. $0 \le x < -\frac{\beta_1}{\sigma^2}$. It implies that $\beta_1 < 0$ and $\exp(\frac{-\beta_1^2}{2\sigma^2}) < \tilde{\mu} \le 1$. Applying numerical methods, we obtain E-optimal designs for some representative values of $\boldsymbol{\beta}$, m and σ^2. The results are listed in Table 2.

Table 2: E-optimal designs for model with random slope.

$\beta_0=1,\beta_1=-5,m=200$			
σ^2	p_1^*	$\tilde{\mu}_1^*$	$\tilde{\mu}_2^*$
0	0.7726	0.0775	1
0.5	0.7493	0.1095	1
1	0.7011	0.1798	1
1.5	0.6684	0.2465	1

As σ^2 increases, optimal point design get more closer to each other. To obtain optimal support points, note that $\log(\tilde{\mu}^*) = \beta_1 x^* + \frac{1}{2}\sigma^2 x^{*2}$ is a quadratic function for the random slope example. By regarding to above assumptions, we have

$$x^* = \frac{-\beta_1 - \sqrt{\beta_1^2 + 2\sigma^2 \log(\tilde{\mu}^*)}}{\sigma^2}.$$

Figure 3 confirms optimality obtained design for special value of parameters. It is showed that the Fréchet derivative is zero at ξ^*.

As model with random intercept, we want to consider efficiency of the standard design $\xi = \left\{ \begin{array}{cc} x_1 & x_2 \\ 0.5 & 0.5 \end{array} \right\}$ with endpoints $x_1 = 0$ and $x_2 = \frac{-\beta_1^2}{\sigma^2}$ using E_{eff}. The results are shown in Table 3.

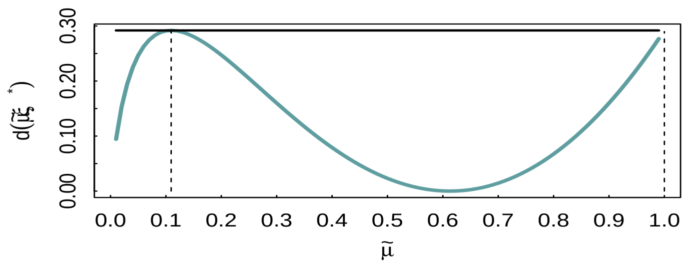

Figure 3: Sensitivity function of ξ^* for $\boldsymbol{\beta}^T = (1, -5)$, $m = 200$ and $\sigma^2 = 0.5$.

Table 3: Efficiency for model with random slope.

σ^2	$\mathrm{E}_{eff}(\beta_0 = -2, \beta_1 = -5, m = 200)$	$\mathrm{E}_{eff}(\beta_0 = 1, \beta_1 = -5, m = 200)$
0	0.082	0.078
0.5	0.099	0.261
1	0.119	0.186
1.5	0.065	0.035

5 Discussion

This article focuses on E-optimal designs in Poisson regression models with random effects. Since the likelihood function has not a closed form, alternately, we use the quasi-likelihood function to obtain the information matrix. This criterion has been discussed in detail for one-variable Poisson regression models in two special situations, random intercept and random slope on specific experimental regions. Also, we discuss the equivalence theorem to confirm obtained optimal designs.

A Proof of Theorem 3.2

First, we make use of the formula $(\boldsymbol{A}_\xi^{-1}+\boldsymbol{C}_\xi)^{-1} = \boldsymbol{A}_\xi-\boldsymbol{A}_\xi(\boldsymbol{I}+\boldsymbol{C}_\xi\boldsymbol{A}_\xi)^{-1}\boldsymbol{C}_\xi\boldsymbol{A}_\xi$, Harville (2013), for the inverse of a sum of matrices. Then we rewrite the quasi-information as $\mathfrak{M}_{\boldsymbol{\beta}}(\xi) = \boldsymbol{F}_\xi^T(\boldsymbol{A}_\xi - \boldsymbol{A}_\xi(\boldsymbol{I} + \boldsymbol{C}_\xi\boldsymbol{A}_\xi)^{-1}\boldsymbol{C}_\xi\boldsymbol{A}_\xi)\boldsymbol{F}_\xi$.

Regarding to this equation, the quasi-information matrix of convex combination of ξ and ξ' for $0 \leq \alpha \leq 1$ is

$$\mathfrak{M}_{\boldsymbol{\beta}}((1-\alpha)\xi + \alpha\xi') = \boldsymbol{F}_{\xi,\xi'}^{\mathrm{T}}[\boldsymbol{A}_{\xi,\xi'}(\alpha)-\boldsymbol{A}_{\xi,\xi'}(\alpha)(\boldsymbol{I}+\boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}_{\xi,\xi'}(\alpha))^{-1}\boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}_{\xi,\xi'}(\alpha)]\boldsymbol{F}_{\xi,\xi'},$$

where $\boldsymbol{F}_{\xi,\xi'} = \begin{pmatrix} \boldsymbol{F}_\xi^T & \boldsymbol{F}_{\xi'}^T \end{pmatrix}^T$ is the joint reduced design matrix for the designs ξ and ξ', $\boldsymbol{A}_{\xi,\xi'}(\alpha) = \begin{pmatrix} (1-\alpha)\boldsymbol{A}_\xi & 0 \\ 0 & \alpha\boldsymbol{A}_{\xi'} \end{pmatrix}$ is the weighted intensity matrix for two designs ξ and ξ' and by $\boldsymbol{C}_{\boldsymbol{\xi},\boldsymbol{\xi'}} = \begin{pmatrix} \boldsymbol{C}_\xi & \boldsymbol{\Gamma}_{\xi,\xi'} \\ \boldsymbol{\Gamma}_{\xi,\xi'}^{\boldsymbol{T}} & \boldsymbol{C}_{\xi'} \end{pmatrix}$ the combined correction matrix, which contains the mixed correction terms for ξ' and ξ in $\boldsymbol{\Gamma}_{\boldsymbol{\xi},\boldsymbol{\xi'}} = (c(x, x'))$, where x and x' are the support points of ξ' and ξ, respectively.

Then for the E-criterion, the directional derivative of $\Psi(\xi)$ in the direction of ξ'

with multiplicity 1 in $\lambda_{max}(\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))$ equals

$$\begin{aligned} F_\Psi(\xi,\xi') &= \frac{d}{d\alpha}\Psi((1-\alpha)\xi+\alpha\xi')|_{\alpha=0^+} \\ &= \frac{d}{d\alpha}\lambda_{max}(\boldsymbol{F}_{\xi,\xi'}^{\mathrm{T}}[\boldsymbol{A}_{\xi,\xi'}(\alpha)-\boldsymbol{A}_{\xi,\xi'}(\alpha)(\boldsymbol{I}+\boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}_{\xi,\xi'}(\alpha))^{-1} \\ &\quad \boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}_{\xi,\xi'}(\alpha)]\boldsymbol{F}_{\xi,\xi'})|_{\alpha=0^+} \\ &= \boldsymbol{q}^T[-\mathfrak{M}_{\boldsymbol{\beta}}^{-1}((1-\alpha)\xi+\alpha\xi')(\boldsymbol{F}_{\xi,\xi'}^{\mathrm{T}}([\boldsymbol{A}'_{\xi,\xi'}(\alpha)-\boldsymbol{A}'_{\xi,\xi'}(\alpha) \\ &\quad (\boldsymbol{I}+\boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}_{\xi,\xi'}(\alpha))^{-1}\boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}_{\xi,\xi'}(\alpha)+\boldsymbol{A}_{\xi,\xi'}(\alpha) \\ &\quad (\boldsymbol{I}+\boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}_{\xi,\xi'}(\alpha))^{-1}\boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}'_{\xi,\xi'}(\alpha)(\boldsymbol{I}+\boldsymbol{C}_{\xi,\xi^*}\boldsymbol{A}_{\xi,\xi'}(\alpha))^{-1} \\ &\quad \boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}_{\xi,\xi'}(\alpha)-\boldsymbol{A}_{\xi,\xi'}(\alpha)(\boldsymbol{I}+\boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}_{\xi,\xi'}(\alpha))^{-1}\boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}'_{\xi,\xi'}(\alpha)] \\ &\quad \boldsymbol{F}_{\xi,\xi'})\mathfrak{M}_{\boldsymbol{\beta}}^{-1}((1-\alpha)\xi+\alpha\xi^*)]\boldsymbol{q}|_{\alpha=0^+}, \end{aligned}$$

where $\mathbf{A}'_{\xi,\xi'}(\alpha)$ is derivative of $\mathbf{A}_{\xi,\xi'}(\alpha)$ w.r.t α.

A simple calculation shows that $(\boldsymbol{I}+\boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}_{\xi,\xi'}(0)) = \begin{pmatrix} \boldsymbol{I}+\boldsymbol{C}_\xi\boldsymbol{A}_\xi & 0 \\ \boldsymbol{\Gamma}_{\xi,\xi'}^{\boldsymbol{T}}\boldsymbol{A}_\xi & \boldsymbol{I} \end{pmatrix}$ which is a lower block triangular with inverse

$$(\boldsymbol{I}+\boldsymbol{C}_{\xi,\xi'}\boldsymbol{A}_{\xi,\xi'}(0))^{-1} = \begin{pmatrix} (\boldsymbol{I}+\boldsymbol{C}_\xi\boldsymbol{A}_\xi)^{-1} & 0 \\ -\boldsymbol{\Gamma}_{\xi,\xi'}^{\boldsymbol{T}}\boldsymbol{A}_\xi(\boldsymbol{I}+\boldsymbol{C}_\xi\boldsymbol{A}_\xi)^{-1} & \boldsymbol{I} \end{pmatrix}.$$

Using multiplication of the block matrix we have

$$\begin{aligned} F_\Psi(\xi,\xi') &= \boldsymbol{q}^T(-\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))[-\boldsymbol{F}_\xi^{\mathrm{T}}\boldsymbol{A}_\xi\boldsymbol{F}_\xi+\boldsymbol{F}_{\xi'}^{\mathrm{T}}\boldsymbol{A}_{\xi'}\boldsymbol{F}_{\xi'}+\boldsymbol{F}_\xi^{\mathrm{T}}\boldsymbol{A}_\xi \\ &\quad (\boldsymbol{C}_\xi\boldsymbol{A}_\xi+\boldsymbol{I})^{-1}\boldsymbol{C}_\xi\boldsymbol{A}_\xi\boldsymbol{F}_\xi+\boldsymbol{F}_{\xi'}^{\mathrm{T}}\boldsymbol{A}_{\xi'}\boldsymbol{\Gamma}_{\xi,\xi'}^{\mathrm{T}}\boldsymbol{A}_\xi(\boldsymbol{C}_\xi\boldsymbol{A}_\xi+\boldsymbol{I})^{-1} \\ &\quad \boldsymbol{C}_\xi\boldsymbol{A}_\xi\boldsymbol{F}_\xi-\boldsymbol{F}_{\xi'}^{\mathrm{T}}\boldsymbol{A}_{\xi'}\boldsymbol{\Gamma}_{\xi,\xi'}^{\mathrm{T}}\boldsymbol{A}_\xi\boldsymbol{F}_\xi-\boldsymbol{F}_\xi^{\mathrm{T}}\boldsymbol{A}_\xi(\boldsymbol{C}_\xi\boldsymbol{A}_\xi+\boldsymbol{I})^{-1} \\ &\quad \boldsymbol{C}_\xi\boldsymbol{A}_\xi(\boldsymbol{C}_\xi\boldsymbol{A}_\xi+\boldsymbol{I})^{-1}\boldsymbol{C}_\xi\boldsymbol{A}_\xi\boldsymbol{F}_\xi-\boldsymbol{F}_\xi^{\mathrm{T}}\boldsymbol{A}_\xi(\boldsymbol{C}_\xi\boldsymbol{A}_\xi+\boldsymbol{I})^{-1} \\ &\quad \boldsymbol{\Gamma}_{\xi,\xi'}\boldsymbol{A}_{\xi'}\boldsymbol{\Gamma}_{\xi,\xi'}^{\mathrm{T}}\boldsymbol{A}_\xi(\boldsymbol{C}_\xi\boldsymbol{A}_\xi+\boldsymbol{I})^{-1}\boldsymbol{C}_\xi\boldsymbol{A}_\xi\boldsymbol{F}_\xi+\boldsymbol{F}_\xi^{\mathrm{T}}\boldsymbol{A}_\xi \\ &\quad (\boldsymbol{C}_\xi\boldsymbol{A}_\xi+\boldsymbol{I})^{-1}\boldsymbol{\Gamma}_{\xi,\xi'}\boldsymbol{A}_{\xi'}\boldsymbol{\Gamma}_{\xi,\xi'}^{\mathrm{T}}\boldsymbol{A}_\xi\boldsymbol{F}_\xi+\boldsymbol{F}_\xi^{\mathrm{T}}\boldsymbol{A}_\xi(\boldsymbol{C}_\xi\boldsymbol{A}_\xi+\boldsymbol{I})^{-1} \\ &\quad \boldsymbol{C}_\xi\boldsymbol{A}_\xi\boldsymbol{F}_\xi-\boldsymbol{F}_\xi^{\mathrm{T}}\boldsymbol{A}_\xi(\boldsymbol{C}_\xi\boldsymbol{A}_\xi+\boldsymbol{I})^{-1}\boldsymbol{\Gamma}_{\xi,\xi'}\boldsymbol{A}_{\xi'}\boldsymbol{F}_{\xi'}](\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))\boldsymbol{q}. \end{aligned}$$

Then, above equation can be represented as

$$\begin{aligned} F_\Psi(\xi,\xi') &= \boldsymbol{q}^T(-\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))[(\boldsymbol{F}_{\xi'}-\boldsymbol{\Gamma}_{\xi,\xi'}^{\mathrm{T}}(\boldsymbol{A}_\xi^{-1}+\boldsymbol{C}_\xi)^{-1}\boldsymbol{F}_\xi)^{\mathrm{T}}\boldsymbol{A}_{\xi'} \\ &\quad (\boldsymbol{F}_{\xi'}-\boldsymbol{\Gamma}_{\xi,\xi'}^{\mathrm{T}}(\boldsymbol{A}_\xi^{-1}+\boldsymbol{C}_\xi)^{-1}\mathbf{F}_\xi)-\boldsymbol{F}_\xi^{\mathrm{T}}(\boldsymbol{A}_\xi^{-1}+\boldsymbol{C}_\xi)^{-1}\boldsymbol{F}_\xi \\ &\quad +\boldsymbol{F}_\xi^{\mathrm{T}}(\boldsymbol{A}_\xi^{-1}+\boldsymbol{C}_\xi)^{-1}\boldsymbol{C}_\xi(\boldsymbol{A}_\xi^{-1}+\boldsymbol{C}_\xi)^{-1}\boldsymbol{F}_\xi)](\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))\boldsymbol{q}. \end{aligned}$$

Niaparast and Schwabe (2013) showed that the directional derivative $F_\Psi(\xi,\xi')$ is linear in ξ'. Therefore it suffices to consider one-point design ξ_x which assign all m observations to one setting x. For such one-point designs the directional derivative reduces to

$$F_\Psi(\xi,\xi_x) = \boldsymbol{q}^T(-\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))[(\boldsymbol{f}(x)-\boldsymbol{F}_\xi^T(\boldsymbol{A}_\xi^{-1}+\boldsymbol{C}_\xi)^{-1}\boldsymbol{c}_{\xi,x})m\mu(x)$$

$$
\begin{aligned}
& (\boldsymbol{f}(x) - \boldsymbol{F}_\xi^T(\boldsymbol{A}_\xi^{-1} + \boldsymbol{C}_\xi)^{-1}\boldsymbol{c}_{\xi,x})^T - \mathbf{F}_\xi^T(\boldsymbol{A}_\xi^{-1} + \boldsymbol{C}_\xi)^{-1}\mathbf{F}_\xi \\
& +\mathbf{F}_\xi^T(\boldsymbol{A}_\xi^{-1} + \boldsymbol{C}_\xi)^{-1}\mathbf{C}_\xi(\boldsymbol{A}_\xi^{-1} + \boldsymbol{C}_\xi)^{-1}\mathbf{F}_\xi)](\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))\boldsymbol{q} \\
= \; & d(x,\xi) - \boldsymbol{q}^T\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi)\boldsymbol{q} + \boldsymbol{q}^T\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi) \\
& [\mathbf{F}_\xi^T(\boldsymbol{A}_\xi^{-1} + \boldsymbol{C}_\xi)^{-1}\mathbf{C}_\xi(\boldsymbol{A}_\xi^{-1} + \boldsymbol{C}_\xi)^{-1}\mathbf{F}_\xi]\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi))\boldsymbol{q}.
\end{aligned}
$$

According to the general equivalence theorem a design ξ^* is optimal, if and only if

$$\forall x \in \mathcal{X};\ F_\Psi(\xi^*, \xi_x) \geq 0.$$

Hence ξ^* is E-optimal if

$$
\begin{aligned}
d(x,\xi^*) \quad \leq \quad & \boldsymbol{q}^T\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*)\boldsymbol{q} - \boldsymbol{q}^T\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*) \\
& [\mathbf{F}_\xi^T(\boldsymbol{A}_{\xi^*}^{-1} + \boldsymbol{C}_{\xi^*})^{-1}\boldsymbol{C}_{\xi^*}(\boldsymbol{A}_{\xi^*}^{-1} + \boldsymbol{C}_{\xi^*})^{-1}\mathbf{F}_{\xi^*}]\mathfrak{M}_{\boldsymbol{\beta}}^{-1}(\xi^*))\boldsymbol{q}.
\end{aligned}
$$

References

Dette, H. and Haines, L.M. (1994). E-optimal designs for linear and nonlinear models with two parameterst. *Biometrika*, **81**:739–754.

Dette, H., Melas, V.B. and Pepelyshev, A. (2004). Optimal designs for a class of nonlinear regression models. *The Annals of Statistics*, **32**(5):2142–2167.

Dette, H. and Studden, W.J. (1993). Geometry of E-optimality. *The Annals of Statistics*, **21**(1):416–433.

Harville, D.A. (1997). *Matrix Algebra From a Statistician's Perspective*. New York: Springer.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *The Annals of Statistics*, **2**(5):849–879.

Niaparast, M. (2009). On optimal design for a Poisson regression model with random intercept. *Statistics & Probability Letters*, **79**(6):741–747.

Niaparast, M. and Schwabe, R. (2013). Optimal design for quasi-likelihood estimation in Poisson regression with random coefficients. *journal of Statistical Planning and Inference*, **143**(2):296–306.

Prus, M. (2019). Various optimality criteria for the prediction of individual response curves. *Statistics & Probability Letters*, **146**:36–41.

Pukelsheim, F. (2006). *Optimal Design of Experiments*. Society for Industrial and Applied Mathematics.

Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. New York: Springer.