Approximate Quasi-Fisher information for designing experiments in mixed ordinal regression models

Parisa Parsamaram

Otto-von-Guericke University Magdeburg, Institute of Mathematical Stochastics

Summary

- D-optimum design for mixed ordinal regression
- Fisher information
- No closed form integral of the marginal likelihood
- Quasi Fisher information related to quasi-likelihood
- Two approximations for the quasi Fisher information
- D-optimum design for the new approximations

Model Specification

- M ordered categories
- $\mathbf{Y}_{ijk} = (Y_{ijk}^{(1)}, Y_{ijk}^{(2)}, ..., Y_{ijk}^{(M)})^\top$
- N units; $i = 1, ..., N$
- t experimental settings per unit; $j = 1, ..., t$
- n_{ij} replication within each experimental setting; $k = 1, ..., n_{ij}$
- Latent utility U_{ijk}: $U_{ijk} = f(\boldsymbol{x}_{ij})^\top \boldsymbol{\beta} + h(\boldsymbol{x}_{ij})^\top \boldsymbol{\zeta}_i + \epsilon_{ijk}$
- $\boldsymbol{\beta}$ fixed effect, $\boldsymbol{\zeta}_i$ individual random effect
- $Y_{ijk}^{(m)} = 1$ if $\gamma_{m-1} \leq U_{ijk} < \gamma_m$, otherwise $Y_{ijk}^{(m)} = 0$
- Thresholds γ_m: $-\infty = \gamma_0 < ... < \gamma_{m-1} < \gamma_m < ... < \gamma_M = \infty$
- $\boldsymbol{x}_{ij}$ is the explanatory variable
- $\delta_j^{(m)}(\boldsymbol{\zeta}_i) = \gamma_m - (f(\boldsymbol{x}_{ij})^\top \boldsymbol{\beta} + h(\boldsymbol{x}_{ij})^\top \boldsymbol{\zeta}_i)$
- $P(Y_{ijk}^{(m)} = 1 \mid \boldsymbol{\zeta}_i) := p_{ij}^{(m)}(\boldsymbol{\zeta}_i)$
- $p_{ij}^{(m)}(\boldsymbol{\zeta}_i) = \Phi(\delta_j^{(m)}(\boldsymbol{\zeta}_i)) - \Phi(\delta_j^{(m-1)}(\boldsymbol{\zeta}_i))$
- $\boldsymbol{\zeta}_i \sim N_q(\mathbf{0}_q, \boldsymbol{\Sigma})$

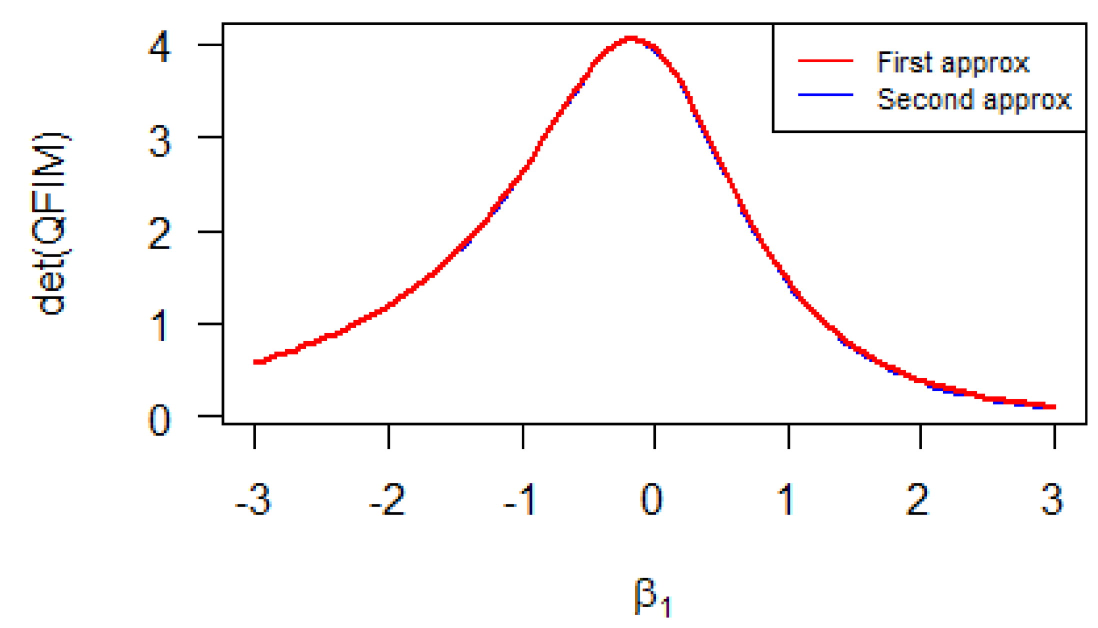
Yatskiv and Kolmakova (2011) investigated the effects of the seven groups of quality particular attributes on the estimates of the overall quality of service of buses and coaches, ranging from 1-5.

Quasi information for approximate design

1. $\xi = \begin{pmatrix} \boldsymbol{x}_1 & \dots & \boldsymbol{x}_t \\ w_1 & \dots & w_t \end{pmatrix}$; $\sum_{j=1}^t w_j = 1$; $w_j = \frac{n_j}{n}$; $w_j \in [0, 1]$.
2. The general formula for the quasi Fisher information matrix is denoted as:
$$M^Q(\xi, \boldsymbol{\beta}) = \mathbf{D}^T \mathbf{V}^{-1} \mathbf{D} \qquad (1)$$
3. $\mathbf{D}$ denotes the derivative of the marginal expectation of the response with respect to $\boldsymbol{\beta}$
4. $\mathbf{V}$ describes the marginal response variance and it depends on $\{w_1, ..., w_t\}$.
For the computation of matrix $\mathbf{V}$ we need to achieve $\mathrm{E}(p_j^{(m)}(\boldsymbol{\zeta}))$ and $\mathrm{E}(p_j^{(m)}(\boldsymbol{\zeta}) p_{j'}^{(m')}(\boldsymbol{\zeta}))$.
5. The first element can be obtained directly by integral solution (Zeger et al. 1988) and the second element is obtained based on the two approximations:
5.1. The numerical computation which is based on the Simpson's rule.
5.2. Solving the second order moment of the integral which is finally obtained as follows:
$$\mathrm{E}(p_j^{(m)}(\boldsymbol{\zeta}_i) p_{j'}^{(m')}(\boldsymbol{\zeta}_i)) = \int_{-\infty}^{\alpha_j^{(m)}} \phi(z) \Phi(q_{j'}^{(m')}(z)) dz;$$
$\alpha_j^{(m)} = \mid \sigma^2 h(\boldsymbol{x}_{ij}) h^\top(\boldsymbol{x}_{ij}) + I_q \mid^{-.5} (\gamma_m - f^\top(\boldsymbol{x}_{ij}) \boldsymbol{\beta})$
and $c_{jj'} = \sigma^2 h_j h_{j'}^\top$, $q_j^{(m)}(z) = \frac{|c_{jj'} + I_q|^{.5}}{|2c_{jj'} + I_q|^{.5}} (\gamma_m - f^\top(\boldsymbol{x}_{ij}) \boldsymbol{\beta} - \frac{\sigma^2 h^\top(\boldsymbol{x}_{ij}) h(\boldsymbol{x}_{ij}) z}{|c_{jj'} + I_q|^{.5}})$.
6. In order to make the matrix V invertible, we need to reduce one level of the response variable Y.

Example: Binary predictor

Random intercept binary model $f(\boldsymbol{x}_{ij}) = (1, \boldsymbol{x}_{ij})^T$, $\boldsymbol{\zeta} = \zeta_0$, $\boldsymbol{\zeta} \sim N(0, \sigma^2)$, $h = 1$, $M = 2$, $x_1 = 0$, $x_2 = 1$:

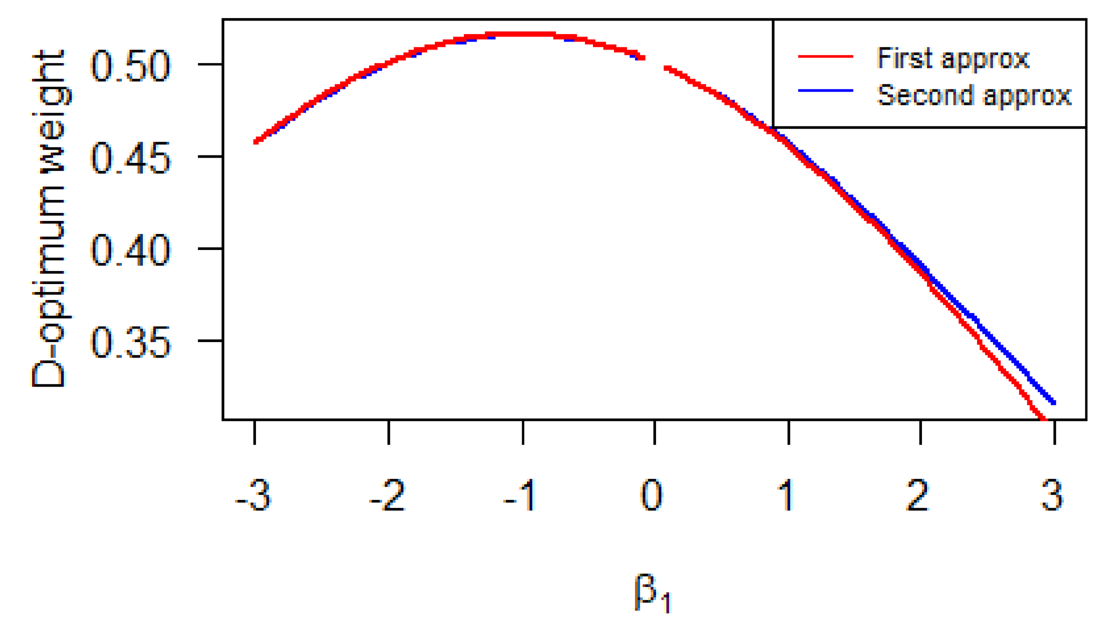


Determinant of the QFIM considering $\beta_1 \in (-3, 3)$, $\beta_0 = 1$, $\sigma^2 = 1$, $n = 20$ because $n_1 = nw^\star, n_2 = n(1 - w_1^\star)$ with $w^\star = .5$

D-optimum Design

Locally D-optimum design with specified experimental settings $x_1 = 0$, $x_2 = 1$:

$$\xi^* = \begin{pmatrix} 0 & 1 \\ w^* & 1 - w^* \end{pmatrix}.$$



D-optimum weight against $\beta_1 \in (-3, 3)$, as $\beta_0 = 1$, $\sigma^2 = 1$, $n = 20$ because $n_1 = nw^\star, n_2 = n(1 - w_1^\star)$ with $w^\star = .5$

For $n = 50$,

$$\xi^{*(1)} = \begin{pmatrix} 0 & 1 \\ 0.52 & 0.48 \end{pmatrix}, \xi^{*(2)} = \begin{pmatrix} 0 & 1 \\ 0.53 & 0.47 \end{pmatrix}. \qquad (2)$$

With respect to two approximations.

Acknowledgement:

I am grateful to my supervisor, professor Rainer Schwabe for his advice and helps in this work.

The work is supported by a grant of the Landesgraduiertenförderung of the state Saxony-Anhalt