学校代码：10270　　分类号：021　　学号：162502921

上海师范大学

硕士专业学位论文

Logistic 回归模型的 R-最优设计

学　　院：数理学院
专业学位类别：应用统计
专业领域：应用统计
研究生姓名：晋珊
指导教师：岳荣先教授
完成日期：2018年05月

论文题目：logistic 回归模型的 R-最优设计

论文类型：理论研究

学科专业：应用统计

学位申请人：晋珊

指导教师：岳荣先

摘要

近些年针对广义线性模型最优设计的研究一直在不断发展。考虑到广义线性模型中 Logistic 回归模型在临床研究、流行病学、预测医学等学术领域的广泛应用，本文分别从固定截距与随机截距 Logistic 回归模型入手，构造了相应模型 R-最优设计的求解方法。

在固定截距 Logistic 回归模型中，考虑单因素与两因素的情形。首先利用方向导数理论，写出具体敏感性函数，推导出 R-最优准则的一般等价性定理，构造出 R-最优设计的迭代算法，然后对单因素模型取初始参数 β_1。给定两个设计点的初始设计，利用迭代算法求解出该设计下的 R-最优设计，并作图检验结果的可靠性与准确性。同样在双因素模型中选取初始参数 β_1 和 β_2，给定九个设计点的初始设计，用相同的方法求解出该设计下的 R-最优设计，并利用相应的图像验证，最终给出固定截距 logistic 回归模型的 R 效率。

在随机截距 Logistic 回归模型中，仅考虑单因素的情形。利用基于 PQL1 估计的参数 β 的协方差矩阵构造准信息矩阵，写出 R-最优准则函数。由于单层随机截距模型等价于固定截距模型，类似固定截距模型的求解方法，对单因素模型取初始参数 β_1，设定随机系数的分布，给定两个设计点的初始设计，利用迭代算法求解该初始设计下的 R-最优设计，并作图来检验结果的可靠性与准确性。

结果表明，固定截距与随机截距 Logistic 回归模型的 R-最优都是存在的，并且计算得到的单因素模型以及双因素模型 R-效率均体现了 R-最优设计相比初始设计而言模型参数效果更好。

关键词：Logistic 回归模型；固定截距；随机截距；R-最优设计

Abstract

In recent years, research on the optimal design of generalized linear models has been continuously developed. Considering that the logistic regression model in the generalized linear model is widely used in the academic fields such as clinical research, epidemiology and predictive medicine, this dissertation starts from the logistic regression model with fixed intercept and random intercept, and constructs the corresponding model R-optimal design Solution method.

In the fixed intercept logistic regression model, consider the single factor and two factors. First, the directional derivative theory is used to write specific sensitivity functions, and the general equivalence theorem of R-optimal criteria is deduced, an iterative algorithm for R-optimal design is constructed, and then the initial parameters β_1 are taken for the single-factor model. Given the initial design of the two design points, an iterative algorithm is used to solve the R-optimal design under the design, and the reliability and accuracy of the verification results are plotted. Also select the initial parameters β_1, β_2 in the two-factor model and give the initial design of nine design points, use the same method to solve the R-optimal design under the design, and use image verification. Finally, the R-efficiency of the fixed-interval logistic regression model is given.

In the random intercept logistic regression model, only the single factor case is considered. The quasi-information matrix is constructed using the covariance matrix based on the parameters of the PQL1 estimation, and the R-optimal criterion function is written. Since the single-layer random intercept model is equivalent to the fixed intercept model, similar to the fixed intercept model, the initial parameters β_1 are set for the single-factor model, the distribution of the random coefficients is set, and the initial design of the two design points is given. An iterative algorithm is used to solve the R-optimal design under the initial design and to verify the reliability and accuracy of the results.

The results show that the R-optimal of fixed-intercept and random intercept

logistic regression models are both present,and the calculated single-factor model and the R-efficiency of the two-factor model all show that the R-optimal design is better than the initial design.

Key words： logistic regression model; fixed intercept; random intercept; R-optimal design

目录

第一章 绪 论

在工农业生产中，常常会遇到如何控制试验次数以节约试验成本或如何搭配工艺参数从而获得最优产量等诸如此类的问题，随着社会生产快速发展的需要，针对如上问题，许多试验设计由此而生，例如旋转设计，正交设计，均匀设计，混料设计等，其中最优试验设计在近四十年发展起来的。它提高了模型的精度，从而使统计分析的性质得到提升。在此期间，人们不断地提出各种最优准则来扩充最优试验设计的内容，目的是为了比较不同设计变量试验的优劣性。目前，已有的最优准则包括 D-最优、A-最优、C-最优、G-最优、I-最优等。

1.1 研究背景

1.1.1 Logistic 回归模型

假设二值变量 y 代表某事件发生的可能性，记 y=1 表示事件发生，y=0 表示事件不发生。如果每次观测有 p 个确定性变量 $x_1, x_2, \cdots, x_p$ 对变量 y 的取值有影响，考虑简单线性回归模型

$$y_i = x_i^T \beta + \varepsilon_i$$

上式 ε_i 称为随机误差项，其均值为 0。若在第 i 次观测中（i =1,…,n ）事件发生的概率和事件不发生的概率分别为

$$P(y_i=1|x_i)=\pi_i=\frac{e^{x_i^T\beta}}{1+e^{x_i^T\beta}}, P(y_i=0|x_i)=1-\pi_i=\frac{1}{1+e^{x_i^T\beta}}$$

根据离散型随机变量期望值的定义，则因变量 y_i 的条件均值可以定义为

$$E(y_i \mid x_i)=\pi_i=\frac{e^{x_i^T\beta}}{1+e^{x_i^T\beta}}$$

由于 y 为取值 0 和 1 的定性变量，为了更准确的描述 Logistic 回归模型，首先引入 Logit 函数，设变量 z 取值范围为[0,1]，则

$$\log it(z)=\ln(\frac{z}{1-z}) \qquad (1\text{-}1)$$

函数（1-1）将变量 z 的取值扩大到 $(-\infty,+\infty)$，这种变换称为 Logit 变换，故 Logistic 回归模型可以表示为

$$\log it(\pi_i)=\ln(\frac{\pi_i}{1-\pi_i})=x_i^T\beta, i=1,\ldots,n. \qquad (1\text{-}2)$$

模型中 $\beta=(\beta_0,\cdots,\beta_p)^T$ 表示回归系数，$x_i=(1,x_{1i},x_{2i},\cdots,x_{pi})^T$ 表示第 i 次观测中的一组自变量。事件发生与不发生的概率比值 $\frac{\pi_i}{1-\pi_i}$ 表示事件的发生比，上述模型也称为 Logistic 固定截距回归模型。

在随机抽样中，数据来源可能具有层次结构，各层次之间保持相互独立，但层次内的观测具有群聚性，故对模型（1-2）中截距项 β_0 取值加入随机效应，即构成新的随机截距 $\beta_{0j}=\beta_0+b_j$，随机截距模型第 j 层（$j=1,2,\cdots,m$）第 i 次观测（$i=1,2,\cdots,n$）的结果表示为 y_{ij}，对于第 j 层，因变量 y_{ij} 的条件期望为

$$E(y_{ij}|x_{ij},b_j)=\pi_{ij}$$

此时 Logistic 随机截距回归模型可以表示为

$$\log it(\pi_{ij})=\ln(\frac{\pi_{ij}}{1-\pi_{ij}})=f_{ij}^T\beta+b_j, i=1,\ldots,n, j=1,\ldots,m. \qquad (1\text{-}3)$$

其中 b_j 代表不同层的随机部分且相互独立，b_j 服从某一分布，模型（1-3）的随机截距部分即为 b_j，$f_{ij}=(1,x_{ij1},\ldots,x_{ijp})^T$，$\beta=(\beta_0,\cdots,\beta_p)^T$，$x_j=x_{\cdot j}=(x_{1j},x_{2j},\ldots,x_{nj})^T$，$X_j=(x_{\cdot j1},x_{\cdot j2},\cdots,x_{\cdot jp})$。

记 $F_j=(1_n\vdots X_j)=(f_{1j},\ldots,f_{nj})^T$，$F_j$ 是 $n\times(p+1)$ 维的矩阵，$Y_j^*=(y_{1j}^*,y_{2j}^*,\ldots,y_{nj}^*)^T$，$\pi_j=(\pi_{1j},\ldots,\pi_{nj})^T$，为了方便描述，随机截距模型第 j 层对应的线性回归模型可定义为

$$Y_j^* \triangleq \log it(\pi_j)=F_j\beta+1_n b_j$$

所有层对应线性回归模型的全体可表示为

$$Y^{*} \triangleq \begin{pmatrix} F_1 \\ \vdots \\ F_m \end{pmatrix} \beta + \begin{pmatrix} 1_n & & 0 \\ & \ddots & \\ 0 & & 1_n \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = F\beta + I_{nm} b. \qquad (1\text{-}4)$$

1.1.2 回归模型的参数估计

在固定截距情况下，对变量 y 进行 n 次观测得到观测值 $y_1, y_2, \cdots, y_n$，显然 $y_i, i = 1, \ldots, n$ 是相互独立的伯努利随机变量，其概率密度函数为

$$p(y_i \mid \pi_i) = \pi_i^{y_i} (1-\pi_i)^{1-y_i}，\ y_i = 0,1$$

可以表示出 y_i 的似然函数

$$L(\pi_i) = \prod_{i=1}^{n} \pi_i^{y_i} (1-\pi_i)^{1-y_i} = \prod_{i=1}^{n} \left(\frac{\pi_i}{1-\pi_i} \right)^{y_i} (1-\pi_i)^{1-y_i}. \qquad (1\text{-}5)$$

带入（1-2）可得其对数似然函数为

$$l(x;\beta) = \sum_{i=1}^{n} \left\{ y_i x_i^T \beta - \ln(1 + \exp(x_i^T \beta)) \right\}.$$

Logistic 固定截距回归模型的极大似然估计需要满足 $\beta = \max\limits_{\beta} l(x;\beta)$，即 $\dfrac{\partial l(x;\beta)}{\partial \beta} = 0$，故

$$\frac{\partial l(x;\beta)}{\partial \beta} \approx \sum_{i=1}^{n} \left[x_i (y_i - \pi_i) \right] = 0.$$

以上称为似然方程组，对于模型（1-2）的参数估计由于其复杂性，一般用迭代法求得似然解，常用的方法有 Newton-Raphson 法。

取 $s_i = y_i - \pi_i$，$v_i = \pi_i(1-\pi_i)$，定义 $S^T = (s_1, \cdots, s_n)$，$V = \mathrm{diag}(v_i)$，对 β 求一、二阶偏导可得

$$\frac{\partial l(\beta)}{\partial \beta} = \sum_{i=1}^{n} x_i (y_i - \pi_i) = X^T S,$$

$$\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta^T} = -\frac{\partial}{\partial \beta^T} \left(\sum_{i=1}^{n} x_i (y_i - \pi_i) \right) = \frac{\partial}{\partial \beta^T} \left(\sum_{i=1}^{n} x_i \pi_i \right) = -\sum_{i=1}^{n} x_i^T \pi_i (1-\pi_i) x = -X^T V X.$$

其中 $X=\begin{pmatrix}1 & x_{11} & \cdots & x_{1p}\\ 1 & x_{21} & \cdots & x_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_{n1} & \cdots & x_{np}\end{pmatrix}$，$S=\begin{pmatrix}y_1-\pi_1\\ y_2-\pi_2\\ \vdots\\ y_n-\pi_n\end{pmatrix}$，$V=\begin{pmatrix}\pi_1(1-\pi_1) & 0 & \cdots & 0\\ 0 & \pi_2(1-\pi_2) & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \pi_n(1-\pi_n)\end{pmatrix}$。

模型参数 β 的牛顿迭代公式为

$$\beta^{(t+1)}=\beta^{(t)}+\left\lfloor (X^TV(\beta^{(t)})X)^{-1}(X^TS(\beta^{(t)}))\right\rfloor. \qquad (1\text{-}6)$$

将上式表示成加权最小二乘估计的形式：

$$\beta^{(t+1)}=(X^TV(\beta^{(t)})X)^{-1}X^TV(\beta^{(t)})Z^{(t)},\ Z^{(t)}=X^T\beta^{(t)}+V(\beta^{(t)})^{-1}S(\beta^{(t)}).$$

当迭代收敛时，参数估计为

$$\hat{\beta}=(X^TV(\hat{\beta})X)^{-1}X^TV(\hat{\beta})Z,\ Z=X^T\hat{\beta}+V(\hat{\beta})^{-1}S(\hat{\beta}). \qquad (1\text{-}7)$$

由于 Rao[35]的研究证明了最大似然估计下参数的协方差阵可以由二阶偏导数矩阵估计出来，即 $Cov(\hat{\beta})=\left(-\frac{\partial^2 l(\beta)}{\partial\beta\partial\beta^T}\right)^{-1}$，故 Logistic 固定截距回归模型的参数协方差阵为

$$Cov(\hat{\beta})=(X^TV(\hat{\beta})X)^{-1}. \qquad (1\text{-}8)$$

在随机截距情况下，给定模型（1-3），可得其响应变量的密度函数为

$$p(y_{ij}|b_j)=\pi_{ij}(b_j)^{y_{ij}}(1-\pi_{ij}(b_j))^{1-y_{ij}},\ y_{ij}=0,1$$

同样带入（1-3）可得似然函数为

$$L(\beta)=\prod_{j=1}^{m}\int_0^{\infty}\prod_{i=1}^{n}\pi_{ij}(b_j)(1-\pi_{ij}(b_j))f(b_j)db_j. \qquad (1\text{-}9)$$

其中 $\pi_{ij}(b_j)=\frac{\exp(f_{ij}^T\beta+b_j)}{1+\exp(f_{ij}^T\beta+b_j)}$，由于上述模型参数的似然函数不是闭合的，故无法迭代求解。

不带截距项分布假设的 logistic 随机截距模型参数 β 的半参数估计量即为下列多项式的根：

$$\sum_{j=1}^{m}Seff(y_j,x_j)=\sum_{j=1}^{m}\sum_{i=2}^{n}(x_{ij}-x_{1j})\left\{Q_{i-1,j}-E(Q_{i-1,j}\mid P_j,x_j)\right\}=0$$

其中$Q_j=(y_{1j},\ldots,y_{nj})^T$，$P_j=\sum_{i=1}^{n}y_{ij}$，而$E(Q_{i-1,j}\mid P_j,X_j)$可以通过$Q_j$的条件概率函数$f(q_j\mid p_j,x_j)$计算获得

$$f(q_j\mid p_j,x_j)=\frac{\exp\left\{\sum_{i=2}^{n}(x_{ij}-x_{1j})^T\beta_{q_{i-1},j}\right\}}{\sum_{\mathrm{B}(q_j)}\exp\left\{\sum_{i=2}^{n}(x_{ij}-x_{1j})^T\beta_{q_{i-1},j}\right\}}.$$

上式中$\mathrm{B}(q_j)$是依赖于p_j，m取值的q_j所有可能的集合，具体构造详见 Garcia 和 Ma[36]，本文不再赘述，并且通过 Abebe[16]的研究可知参数β基于 PQL1(First order penalized quasi-likelihood)估计的协方差矩阵为

$$Cov(\hat{\beta})=(F^TV^{-1}F)^{-1},\qquad(1\text{-}10)$$

记$V=diag(V_1,\ldots,V_m)$，V_j与第j层的协方差矩阵有关：

$$V_j\approx Var(y_j\mid b_j)^{-1}+1_jD1_j^T,\quad Var(y_j\mid b_j)=diag(Var(y_{1j}\mid b_j),\ldots,Var(y_{nj}\mid b_j)).$$

D表示随机项b_j的协方差矩阵，假设b_j服从$N(0,\sigma_b{}^2)$，则$D=\sigma_b{}^2$。

已知 Logistic 随机截距模型中观测y_{ij}的条件期望和方差为

$$\begin{aligned}&E(y_{ij}\mid b_j)=\pi_{ij}(b_j),Var(y_{ij}\mid b_j)=\pi_{ij}(b_j)(1-\pi_{ij}(b_j)),\\&Var(y_{ij})=E(Var(y_{ij}\mid b_j))+Var(E(y_{ij}\mid b_j))\\&\qquad=E(\pi_{ij}(b_j)(1-\pi_{ij}(b_j)))+Var(\pi_{ij}\mid b_j).\end{aligned}$$

可得

$$V_j=\begin{pmatrix}(\pi_{1j}(1-\pi_{1j}))^{-1}+\sigma_b{}^2 & \sigma_b{}^2 & \cdots & \sigma_b{}^2\\ \sigma_b{}^2 & (\pi_{2j}(1-\pi_{2j}))^{-1}+\sigma_b{}^2 & \cdots & \sigma_b{}^2\\ \cdots & \cdots & \ddots & \vdots\\ \sigma_b{}^2 & \sigma_b{}^2 & \cdots & (\pi_{nj}(1-\pi_{nj}))^{-1}+\sigma_b{}^2\end{pmatrix}_{n\times n}.\qquad(1\text{-}11)$$

由于整体观测Y中不同层之间是相互独立的，故变量Y的协方差阵是一个对角阵，即$W=diag(W_1,\ldots,W_m)$，W_j表示第j层的协方差矩阵。根据 Ouwens 和 Tan[23]的研究可知随机截距模型（1-4）参数估计的渐进协方差阵为

$$Cov(\hat{\beta}) \approx \left(\sum_{j=1}^{m} \left(\frac{\partial \pi_j(\beta)}{\partial \beta} \right)^T W_j^{-1} \frac{\partial \pi_j(\beta)}{\partial \beta} \right)^{-1} = (\sum_{j=1}^{m} F_j^T W_j F_j)^{-1}, \quad (1\text{-}12)$$

$$W_j \approx Var(y_j | b_j = 0) + Var(y_j | b_j = 0) 1_j D 1_j^T Var(y_j | b_j = 0)$$

本文的研究将围绕基于 PQL1 估计的协方差矩阵进行。

1.2 研究的意义与目的

Logistic 回归模型作为广义线性模型中比较常见的一种，属于概率型非线性回归模型，它是研究二分类观察结果与一些影响因素之间关系的一种多变量分析方法，由于其响应变量的特殊性和处理分类数据的有效性，故常被用于医学、经济学、生物学、犯罪心理学、工程技术学等领域。传统的 Logistic 回归模型多为固定截距模型，但是在研究重复测量或具有层次结构的数据时，层次内部的数据会具有群聚性，例如医学领域的横断面调查数据，此时固定截距模型对数据的分析会不准确，而 Logistic 随机截距回归模型得到研究结果会更合理。有关 Logistic 回归模型的 R-最优设计可以帮助人们在已知真实模型为 Logistic 回归模型的先验知识下，更好的选择设计点进行试验设计，从而获得理想的模型参数结果。

本文是建立 Logistic 固定截距和随机截距模型基础上 R-最优试验设计，整体论文需要覆盖一下目标：

（1）Logistic 固定截距回归模型的 R-最优设计一般等价性定理的构造，写出敏感性函数，在确定初始参数与设计的情况下，利用迭代算法，针对不同变量个数的模型求解出 R-最优设计的数值解，画出对应图像验证。

（2）考虑更一般化的模型结构，研究带随机截距项的 Logistic 回归模型的 R-最优设计准则的构造，给定初始参数与设计的条件下，利用迭代算法，针对单因素的情况下求解出简单模型的 R-最优设计数值解，画出对应图像验证。

1.3 本文的主要内容

本文主要研究固定截距和随机截距情况下 Logistic 模型的 R-最优设计，分别构造单变量和两变量的固定截距模型以及单变量单层随机截距模型在固定设计

域下的 R-最优准则，同时给出了固定截距模型下 R-最优设计的一般等价性定理和求解算法，并采用迭代算法求出最优结果，最后举例数值模拟出不同参数假设下的 R-最优设计，计算出 R-效率。本文主要由五个部分组成。

第一章介绍了 Logistic 固定截距模型以及随机截距模型的一些简单性质和参数估计，同时说明了本文的研究目的意义和目前国内外的研究情况。

第二章介绍了构造 R-最优设计所需要的铺垫知识，包含信息矩阵、常用设计准则、固定截距模型一般等价性定理的推导和设计不变性理论的说明。

第三章构造了固定截距 Logistic 回归模型在单变量和两变量情况下的 R-最优设计。首先写出相应的模型结构，获得信息矩阵后根据一般等价性定理推导出单变量固定截距 Logistic 回归模型的敏感性函数，同时给出了 R-最优求解的迭代算法，分别给定单变量情况和双变量情况下 β 的参数值范围，列表得出单因素模型和双因素模型在初始设计下的 R-最优设计，并给出图像验证，最后分别计算出 R-效率。

第四章构造了随机截距 Logistic 回归模型在单层单变量情况下的 R-最优设计。先写出具体的随机模型结构，得到 R-最优设计的准则函数，类似固定截距情况一般等价性定理，给定 β_1 的参数值范围，利用迭代算法获得单层模型的 R-最优设计，最后给出图像加以验证。

第五章是对全文所给信息进行总结，指出不足与待深入解决的问题。

第二章 文献综述与相关理论

2.1 文献综述与现状分析

在 20 世纪 20 年代，Fisher 开创性的在农业生产中应用试验设计的统计方法，并出版著作阐述了试验设计的方法论。最优试验设计理论最早是由 Smith 在 1918 年研究多项式模型中提出来的。20 世纪 70 年代初，Fedorov[1],[2]和 Kiefer[3]的研究组成了最优设计的理论核心。

在统计学的讨论中，非线性模型常用在医药学领域，但是非线性模型由于本身的复杂性，在最优设计这方面的研究相对较少。在进行最优设计时，选取不同的最优准则，某一试验估计的效率会因此产生差异。现已被人们所熟知的最优准则包括 D-最优准则和 A-最优准则等。近三十年内，经过多位科学家的不断努力，最优试验设计的内容变得愈加丰富，其中人们讨论内容最多最完整的便是 D-最优准则设计。虽然 D-最优准则的理论与应用都已经比较充分，但是涉及高维度的模型其计算相关的密集椭球体并不容易解出，因此 Holger[4]提出了由 D-最优准则的最小化参数置信椭球体积转换成求解基于 Bonferroni t-区间法的最小化置信矩形体积的 R-最优准则。R-最优准则不仅具有统计学的解释意义，而且满足设计不变性，即对导出的设计空间进行非奇异线性变化，原设计的最优性仍然保持，这是一个非常有用的属性。有关 R-最优设计的在国内应用，赵洪雅和关颖男等[5]研究了二阶可加混料模型的 R-最优设计；孙超[6]讨论了随机系数回归模型的 I_L-最优和 R-最优设计；岳荣先和刘欣[7],[8],[9],[10]则已经从多响应模型和随机系数回归模型方面进行了研究；徐靖[11]研究了二次响应模型的 R-最优设计并比较了最优设计的效率问题。

早在 1838 年，Verhuist 就首次提出了 Logistic 回归模型，并将其应用在人口统计学中的增长曲线中。20 世纪 60 年代初，Cornfield 研究出适合临床研究的 Logistic 回归模型。如今 logistic 回归模型在发达国家被广泛的应用于临床研究、流行病学、预测医学等学术领域。针对 Logistic 回归模型的最优设计，Joy 和 Wong[12]计算了 Logistic 回归的单个变量在立方体设计空间的 D-最优设计；Torsney 和 Gunduz[13]提出了高维 Logistic 回归模型的最优设计；Linda 等[14]研究了 logistic 模型两个变量的局部 D-最优设计；Habib[15]讨论了关于 Logistic 模型

三个独立变量的 D-最优设计；Abebe 等[16]研究了自相关 Logistic 混合模型的贝叶斯 D-最优设计的。而在国内，针对该模型的研究主要集中在某一类数据的应用方面，例如李长平等[17]、陈晓兰和任萍[18]以及张乐勤和陈发奎[19]都采用 Logistic 回归模型来对某一方面进行分析。

特别的，针对同一试验个体在某一项指标上的重复测量问题。一般经典的回归模型假设下观测值之间是相互独立的，而线性混合模型则对变量的连续性有要求。针对离散型或者计数型变量，广义线性混合模型由此提出，通过在模型中纳入随机效应，很好的解释了数据之间的分层随机性。在国外，很多学者将最优设计应用到具有随机效应的混合模型中：Cheng[20]研究了随机区组效应模型下的最优设计；Silvio 和 Anthony[21]研究了多元 Logistic 模型的贝叶斯最优设计和 D-最优设计；Tan 和 Berger[22]研究了线性随机效应模型的最优设计；Ouwens[23]和 Tekle[24]分别研究了 Logistic 混合效应模型的极小化极大准则的最优设计；Debusho 和 Haines[25 讨论了离散设计域上线性随机截距模型的 V-最优和 D-最优；Tommasi 等[26]研究了 Logistic 随机效应模型的 D-最优和 A-最优设计，Habib[27]研究了 Logistic 随机截距模型的贝叶斯 D-最优设计。国内方面，程靖和岳荣先[28],[29]分别研究了恒等设计假设下两变量随机系数模型的最优设计；周晓东[30]研究了线性混合效应模型的最优设计；程靖和岳荣先等[31]研究了带有异方差性的随机系数回归模型的最优设计。

其他有关复杂 Logistic 模型或最优设计的研究还有 Li[32]研究了四个变量的 Logistic 回归模型的 D-最优设计；Adewale[33]以及 Mok[34]研究了误设条件下 Logistic 模型的稳健设计。本文主要考虑 Logistic 固定截距与随机截距模型的简单情况，寻找构造该类模型的 R-最优设计。

2.2 相关理论

2.2.1 信息矩阵

在设计区域 χ 上，$X_1, X_2, \cdots X_n$ 是 χ 中 n 个试验点，n 为一个正整数，则设计 $\xi=\{X_1, X_2, \cdots X_n\}$ 称为一个精确设计。若试验总次数为 N，此时每个试验点可能存在重复试验情况，且重复次数假设为 $r_1, \cdots, r_n$，$r_1+r_2+\cdots+r_n=N$，当总试验

次数 N 变化时，满足以上设计 ξ 的设计点，其重复实验次数不一定仍是正整数，故可引入近似设计的概念。

近似设计 ξ 的分布可以假设如下表示

$$\xi=\begin{Bmatrix} x_{(1)} & \cdots & x_{(n)} \\ w_1 & \cdots & w_n \end{Bmatrix}, x_{(n)}=(x_{1n},\cdots,x_{pn}), 0<w_i<1, \sum_{i=1}^{n} w_i=1,$$

其中 w_n 为单个试验点在整个试验中发生的概率，即 $w_i=\frac{r_i}{N}$， $r_i, i=1,\cdots,n$ 为每个试验点的重复次数，$\xi\in\Xi$，Ξ 为定义在设计区域 χ 上近似设计的全体。

记 $f(x)=(1,x_1,\ldots,x_p)^T$，则该设计的信息矩阵可以定义为

$$M(\beta;\xi)=\sum_{i}^{n} w_i f(x) f(x)^T=\int_{\chi} f(x) f(x)^T d\xi.$$

信息矩阵也称 Fish 信息矩阵既包含了设计点的信息也包含了假设回归模型的信息, 一般情况下，信息矩阵为最小二乘估计方差的逆矩阵。

在确定 Logistic 回归模型的信息矩阵之前，先确定参数估计的方差，以固定截距模型（1-2）为例，根据 1.2 节可知 Logistic 回归模型中 β 的参数估计以及估计的方差，由 Atkinson 和 Donev[37]的研究可知，广义线性模型在设计 ξ 下的信息矩阵为

$$M(\beta,\xi)=\int_{x\in\chi} v(x_1,\ldots,x_p) f(x_1,\ldots,x_p) f(x_1,\ldots,x_p)^T d\xi, \qquad (2\text{-}1)$$

在固定截距条件下，记 $v(x_1,\ldots,x_{p-1})=V^{-1}(\pi)\left(\frac{d\pi}{d\eta}\right)^2$，连接函数 $\eta=\ln(\frac{\pi}{1-\pi})$，$V(\pi)$ 为响应变量 y 的方差函数，并且

$$V^{-1}(\pi)=\frac{1}{\pi(1-\pi)}, \frac{d\eta}{d\pi}=\frac{1}{\pi(1-\pi)},$$

可得

$$v(x_1,\ldots,x_p)=\frac{1}{\pi(1-\pi)}\cdot\left[\pi(1-\pi)\right]^2=\pi(1-\pi). \qquad (2\text{-}2)$$

所以对于 Logistic 固定截距回归模型，近似设计 ξ 的信息矩阵可以写为

$$M(\beta;\xi)=\sum_{i}^{n} w_i M(\beta;x_i)=\int_{\chi} v_i f(x) f(x)^T d\xi. \qquad (2\text{-}3)$$

其中 $v_i=\pi_i(1-\pi_i)$，显然 $Var(\hat{\beta})=M(\hat{\beta};x_i)^{-1}$。

同样，以随机截距模型（1-4）为例，不妨假设每层的设计相同，即只考虑恒等设计，则第 j 层的近似设计 ξ^* 的分布为

$$\xi^*=\begin{Bmatrix} X_{(1j)} & \cdots & X_{(nj)} \\ w_{1j} & \cdots & w_{nj} \end{Bmatrix}, X_{(ij)}=(x_{ij1},\cdots,x_{ijp}), w_{ij}=\frac{1}{n}=w, \sum_{i=1}^{n} w_{ij}=1,$$

假设全体的近似设计为 ξ，$\xi=(\xi^*,\ldots,\xi^*)$，ξ 的概率分布为

$$\xi=\begin{Bmatrix} \xi^* \\ 1 \end{Bmatrix}$$

由于待估参数的似然函数不闭合导致无法获得 Fish 信息矩阵，所以构造 Quasi-likelihood 准信息矩阵来代替 Fish 信息矩阵，记 $F_j=(f_{1j},\ldots,f_{nj})^T$，$f_{ij}=(1,x_{ij1},\ldots,x_{ijp})^T$，根据 Habib[27]的和 Abebe[16]研究可知准信息矩阵表示为：

$$M(\beta;F_j)=Cov(\hat{\beta})_j^{-1}=F_j^T V_j^{-1} F_j. \qquad (2\text{-}4)$$

其中

$$V_j=\begin{pmatrix} (\pi_{1j}(1-\pi_{1j}))^{-1}+\sigma_b^{\,2} & \sigma_b^{\,2} & \cdots & \sigma_b^{\,2} \\ \sigma_b^{\,2} & (\pi_{2j}(1-\pi_{2j}))^{-1}+\sigma_b^{\,2} & \cdots & \sigma_b^{\,2} \\ \cdots & \cdots & \ddots & \vdots \\ \sigma_b^{\,2} & \sigma_b^{\,2} & \cdots & (\pi_{nj}(1-\pi_{nj}))^{-1}+\sigma_b^{\,2} \end{pmatrix}.$$

故对于 Logistic 随机截距模型，近似设计 ξ^* 的准信息矩阵为：

$$\begin{aligned} M_j(\beta;\xi^*)&=\sum_{i}^{n} w_{ij} M(\beta;X_j), \\ M(\beta;\xi^*)&=\sum_{j}^{m} M_j(\beta;\xi^*)=\sum_{j}^{m}\sum_{i}^{n} w_{ij} M(\beta;X_j)=\int_{\chi} F_j^T V_j^{-1} F_j d\xi^*. \end{aligned} \qquad (2\text{-}5)$$

2.2.2 最优设计准则

所谓最优设计是某一设计准则下待估参数的最优估计，而设计准则是基于设计区域 χ 上的设计点 $X_1,X_2,\cdots X_n$ 与其对应的观测值 $Y_1,Y_2,\cdots,Y_n$ 所构造回归模型

中预测响应与响应均值之间拟合程度的函数。由于人们追求实际生产试验的最优化，使得最优设计理论研究发展迅速，最优准则的内容也不断扩充，根据信息矩阵准则函数的不同，目前最常见的最优准则有四种：

（1）D-最优准则

准则函数为

$$\psi_D(\xi)=—\ln\det M(\xi). \qquad (2\text{-}6)$$

若存在设计 $\xi^* \in \Xi$，使得 $\psi_D(\xi^*)=\min\psi_D(\xi)$，则设计 ξ^* 为 D-最优设计，设计的信息矩阵行列式为 $\det M(\xi)$，不难发现，设计达到 D-最优即使信息矩阵的行列式达到最大。

（2）A-最优准则

准则函数为

$$\psi_A(\xi)=\begin{cases} tr\left\lfloor M^{-1}(\xi)\right\rfloor, & 当 \det M(\xi)\neq 0 \\ \infty, & 当 \det M(\xi)=0 \end{cases}. \qquad (2\text{-}7)$$

若存在设计 $\xi^* \in \Xi$，使得 $\psi_A(\xi^*)=\min\psi_A(\xi)$，则设计 ξ^* 为 A-最优设计，tr 表示矩阵的迹，$tr\left\lfloor M^{-1}(\xi)\right\rfloor$ 表示参数向量各分量估计的差异之和，某一设计达到 A-最优即使 $tr\left\lfloor M^{-1}(\xi)\right\rfloor$ 达到最小。

（3）G-最优准则

准则函数为

$$\psi_G(\xi)=\begin{cases} \max d(x,\xi), & 当 \det M(\xi)\neq 0 \\ \infty, & 当 \det M(\xi)=0 \end{cases}. \qquad (2\text{-}8)$$

若存在设计 $\xi^* \in \Xi$，使得 $\psi_G(\xi^*)=\min\psi_G(\xi)$，则设计 ξ^* 为 G-最优设计，$d(x,\xi)$ 是设计的标准化方差函数，且 $d(x,\xi)=f(x)^T M(\xi)^{-1} f(x)$，故设计达到 G-最优即使设计的预测方差在设计空间内的最大值最小。

（4）E-最优准则

准则函数为

$$\psi_E(\xi)=\begin{cases} \lambda_\xi^{-1}, & 当 \det M(\xi)\neq 0 \\ \infty, & 当 \det M(\xi)=0 \end{cases}. \qquad (2\text{-}9)$$

若存在设计$\xi^* \in \Xi$，使得$\psi_E(\xi^*)=\min\psi_E(\xi)$，则设计$\xi^*$为 E-最优设计，$\lambda_{\xi}^{-1}$是设计$\xi$的信息矩阵逆矩阵的最大特征值，故设计达到 E-最优即使信息矩阵逆矩阵的最大特征值达到最小。

以上最优设计中，D-最优设计是研究内容最多讨论最频繁的，虽然 D-最优准则的应用已经比较完善，但是涉及高纬度模型的密集椭球体并不容易解出，因此 Dette[4]的研究中提出了由 D-最优准则的最小化参数置信椭球体积转换成求解基于 Bonferroni t-区间法的最小化置信矩形体积的 R-最优准则，简单来说 R-最优即使得设计信息矩阵的对角元素乘积值达最小。

2.2.3 一般等价性定理

为了求解设计的 R-最优性，可以应用 R-最优设计的一般等价性定理，本节主要推导出固定截距模型 R-最优设计的一般等价性定理及其所具有的设计不变性证明。

定义 2.1 对于 Logistic 固定截距回归模型，定义一个设计$\xi^* \in \Xi$，假设$M(\xi)$可逆，若其可以使准则函数

$$\psi(M(\xi))=\prod_{i=1}^{p}(M^{-1}(\xi))_{ii}=\prod_{i=1}^{p}e_i^T M^{-1}(\xi)e_i. \qquad (2\text{-}10)$$

达到最小值，其中e_i^T为$\mathbb{R}$中单位矩阵的第i个列向量，设计ξ^*存在于Ξ，那么设计ξ^*就是关于未知参数的 R-最优设计。

接下来说明 R-准则的凸性并证明其等价的方向导数：

引理 2.1 对于任意的设计$\xi,\bar{\xi} \in \Xi$，且$0<\alpha<1$，假设$\xi_\alpha=(1-\alpha)\xi+\alpha\bar{\xi}$，则$M(\xi_\alpha)=M((1-\alpha)\xi+\alpha\bar{\xi})$。由于$M(\xi),M(\bar{\xi})$是正定的，故$M(\xi_\alpha)$也是正定的。针对 Logistic 回归模型的 R-最优准则$\psi(M(\xi_\alpha))$，是$M(\Xi)$上的凸函数，即

$$\psi(M(\xi_\alpha))=\psi(M((1-\alpha)\xi+\alpha\bar{\xi}))\le(1-\alpha)\psi(M(\xi))+\alpha\psi(M(\bar{\xi})) \qquad (2\text{-}11)$$

且其在ξ处沿着$\bar{\xi}$方向的 Frechet 导数可以表示为

$$F_{\phi}(\xi,\bar{\xi})=\psi(M(\xi))\left(p-tr\left\{(M^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi))\sum_{i=1}^{p}\frac{e_ie_i^T}{e_i^TM^{-1}(\xi)e_i}\right\}\right).\quad(2\text{-}12)$$

证明：由以上条件可知

$$\begin{aligned}M(\xi_\alpha)&=\int_{\chi}v_i f(x)f(x)^T d\xi_\alpha=(1-\alpha)\int_{\chi}v_i f(x)f(x)^T d\xi+\alpha\int_{\chi}v_i f(x)f(x)^T d\bar{\xi}\\&=(1-\alpha)M(\xi)+\alpha M(\bar{\xi}),\end{aligned}$$

$$\begin{aligned}\frac{dM^{-1}(\xi_\alpha)}{d\alpha}&=-(M^{-1}(\xi_\alpha))^2\frac{d}{d\alpha}M(\xi_\alpha)=-(M^{-1}(\xi_\alpha))^2(M(\bar{\xi})-M(\xi))\\&=(M^{-1}(\xi_\alpha))^2(M(\xi)-M(\bar{\xi})),\end{aligned}$$

$$\begin{aligned}\frac{d\psi(M(\xi_\alpha))}{d\alpha}&=\frac{d}{d\alpha}\prod_{i=1}^{p}e_i^TM^{-1}(\xi_\alpha)e_i=\sum_{i=1}^{p}\left(\prod_{\substack{j=1\\i\neq j}}^{p}e_j^TM^{-1}(\xi_\alpha)e_j\right)\frac{d}{d\alpha}(e_i^TM^{-1}(\xi_\alpha)e_i)\\&=\sum_{i=1}^{p}\left(\prod_{\substack{j=1\\i\neq j}}^{p}e_j^TM^{-1}(\xi_\alpha)e_j\right)e_i^T(M^{-1}(\xi_\alpha))^2(M(\xi)-M(\bar{\xi}))e_i,\end{aligned}$$

$$\begin{aligned}F_{\phi}(\xi,\bar{\xi})&=\lim_{\alpha\to0^+}\frac{d\psi(M(\xi_\alpha))}{d\alpha}\\&=\sum_{i=1}^{p}\left(\prod_{\substack{j=1\\i\neq j}}^{p}e_j^TM^{-1}(\xi)e_j\right)e_i^T(M^{-1}(\xi)-M^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi))e_i\\&=\sum_{i=1}^{p}\frac{\psi(M(\xi))}{e_i^TM^{-1}(\xi)e_i}(e_i^T(M^{-1}(\xi)-M^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi))e_i)\\&=\psi(M(\xi))\sum_{i=1}^{p}\frac{e_i^T(M^{-1}(\xi)-M^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi))e_i}{e_i^TM^{-1}(\xi)e_i}\\&=\psi(M(\xi))\sum_{i=1}^{p}\left(1-\frac{e_i^T(M^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi))e_i}{e_i^TM^{-1}(\xi)e_i}\right)\\&=\psi(M(\xi))\left(p-\sum_{i=1}^{p}\frac{e_i^T(M^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi))e_i}{e_i^TM^{-1}(\xi)e_i}\right)\\&=\psi(M(\xi))\left(p-tr\left\{M^{-1}(\xi)M(\bar{\xi})M^{-1}(\xi)\sum_{i=1}^{p}\frac{e_i^Te_i}{e_i^TM^{-1}(\xi)e_i}\right\}\right).\end{aligned}$$

记δ_x为在设计点x上的设计，其权重为1，δ_x可以表示为$\delta_x=\begin{Bmatrix}x\\1\end{Bmatrix}$，当$\psi(\xi)$中设计$\xi$在设计域$\chi$上任一设计点都可微时

$$F_{\phi}(\xi,\bar{\xi})=\sum_{x}w(x)F_{\phi}(\xi,\delta_x).$$

上式$w(x)$表示设计$\bar{\xi}$在点x处的权重，具体证明可以参考 Liu[8]和方开泰[37]的研究，并且我们称$\phi(x,\xi)=F_{\phi}(\xi,\delta_x)$为敏感性函数。

定义 2.2 本文模型的敏感性函数为

$$\phi(x,\xi)=tr\left\{v(x)M^{-1}(\xi)f^{T}(x)f(x)M^{-1}(\xi)\sum_{i=1}^{p}\frac{e_i^{T}e_i}{e_i^{T}M^{-1}(\xi)e_i}\right\}. \qquad (2\text{-}13)$$

故由定义 2.1、定义 2.2 以及引理 2.1，我们可以得到 Logistic 固定截距回归模型的 R-最优设计的一般等价性定理。

定理 2.1 已知 R-最优准则$\psi(M(\xi))$，$\phi(x,\xi)$为其敏感性函数，若ξ^{*}为 R-最优设计，则$\phi(x,\xi)$满足以下条件

（i）ξ^{*}使得准则$\psi(M(\xi))$取值最小。

（ii）对于任意设计$\xi\in\Xi$，都有$\inf\limits_{x\in\chi}\phi(x,\xi)\le\inf\limits_{x\in\chi}\phi(x,\xi^{*})$。

（iii）$\min\limits_{x\in\chi}\phi(x,\xi^{*})=0$，且最小值在$\xi^{*}$在支撑点上取得。

定理 2.2 若$M(\xi)$是非奇异的

$$\phi(x,\xi)=tr\left\{\sum_{i=1}^{p}\frac{v_i(e_i^{T}M^{-1}(\xi)f(x))^{2}}{e_i^{T}M^{-1}(\xi)e_i}\right\}\le p. \qquad (2\text{-}14)$$

其中$v_i=\pi(1-\pi)$，对于$\forall x\in\chi$，若$\sup\limits_{x\in\chi}\phi(x,\xi^{*})=p$，则$\xi^{*}\in\Xi$为一个 R-最优设计。

值得一提的是，R-最优设计所具有的设计不变性可以简化设计空间变换的计算问题，所谓不变性即对给定的设计空间进行非奇异线性变化到某一新的设计空间，原最优设计ξ^{*}的驻点和最优性仍然保持不变。

证明：给定原始设计空间χ，在此空间下的最优设计为ξ^{*}，其柱点为x，对设计空间χ进行非奇异变换Q至新设计空间$\aleph$，且$z=Qf(x)$，此时信息矩阵为

$$M(z)=\int_{\aleph}f(z)f^{T}(z)d\zeta(z)=Q\int_{\chi}f(x)f^{T}(x)d\zeta(x)Q^{T},$$

对于$\forall z\in\aleph$，

$$\frac{v_i(e_i^T M^{-1}(z) f(z))^2}{e_i^T M^{-1}(z) e_i} = \frac{v_i\left(e_i^T\left(\int_{\aleph} f(z) f^T(z) d\zeta(z)\right)^{-1} f(z)\right)^2}{e_i^T\left(\int_{\aleph} f(z) f^T(z) d\zeta(z)\right)^{-1} e_i}$$

$$= \frac{v_i\left(e_i^T\left(Q\int_{\chi} f(x) f^T(x) d\zeta(x) Q^T\right)^{-1} Q f(x)\right)^2}{e_i^T\left(Q\int_{\chi} f(x) f^T(x) d\zeta(x) Q^T\right)^{-1} e_i}$$

$$= \frac{v_i\left(e_i^T Q^{-1}\left(\int_{\chi} f(x) f^T(x) d\zeta(x)\right)^{-1} (Q^T)^{-1} Q f(x)\right)^2}{Q^{-1} e_i^T\left(\int_{\chi} f(x) f^T(x) d\zeta(x)\right)^{-1} e_i (Q^T)^{-1}}$$

$$= \frac{v_i(e_i^T M^{-1}(x) f(x))^2}{e_i^T M^{-1}(x) e_i}.$$

根据定理 2.2 可知满足本文模型 R-最优设计的充要条件为 $\phi(x,\xi) \le p$，即 $\phi(z,\xi) \le p$，故 z 为设计空间 $\aleph$ 的柱点且该点下的设计最优性保持不变。

第三章 固定截距的 Logistic 回归模型

3.1 固定截距模型及信息矩阵

3.1.1 单变量模型

假设 χ 为设计空间，ε_i 是随机误差，并且服从零均值，同方差，相互独立的条件，当某次观测试验点为 $X=(x_{(1)},\cdots,x_{(n)})$ 时，定义连接函数，构造固定截距的单变量 Logistic 回归模型，响应变量 y 满足二项分布：

$$E(y_i|x_i)=\pi_i=\frac{e^{x_i^T\beta}}{1+e^{x_i^T\beta}},\log it(\pi_i)=\ln(\frac{\pi_i}{1-\pi_i})=f(x)^T\beta=x_i^T\beta \quad (3\text{-}1)$$

其中 $\pi_i=P(y_i=1|x_i)=\frac{e^{x_i^T\beta}}{1+e^{x_i^T\beta}}, x_i^T=(1,x_{1i}), \beta^T=(\beta_0,\beta_1)$。

单变量固定截距 Logistic 回归模型的信息矩阵可以表示为

$$M(\beta;\xi)=\sum_i^n w_i v_i f(x)f(x)^T=\begin{pmatrix}\sum_{i=1}^n w_i\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2} & \sum_{i=1}^n w_i\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2}x_{1i}\\ \sum_{i=1}^n w_ix_{1i}\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2} & \sum_{i=1}^n w_ix_{1i}\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2}x_{1i}\end{pmatrix} \quad (3\text{-}2)$$

其中 $f(x)=(1,x_1), v_i=\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2}$。

以上模型信息矩阵的逆矩阵为

$$M^{-1}(\beta;x_i)=\frac{\begin{pmatrix}\sum_{i=1}^n w_ix_{1i}\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2}x_{1i} & -\sum_{i=1}^n w_i\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2}x_{1i}\\ -\sum_{i=1}^n w_ix_{1i}\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2} & \sum_{i=1}^n w_i\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2}\end{pmatrix}}{\sum_{i=1}^n w_i\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2}\cdot\sum_{i=1}^n x_{1i}w_i\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2}x_{1i}-\left(\sum_{i=1}^n w_i\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2}x_{1i}\right)^2} \quad (3\text{-}3)$$

由于信息矩阵中包含有 $\frac{e^{\beta_0+\beta_1x_{1i}}}{(1+e^{\beta_0+\beta_1x_{1i}})^2}$，故信息矩阵的取值与待估参数向量 β 的选取有关，所以在确定信息矩阵与准则之前，本文预先定义 β 的取值范围。

3.1.2 两变量模型

前文详细介绍了单变量的固定截距 Logistic 回归模型，我们可以写出两变量的固定截距 Logistic 回归模型

$$\log it(\pi_i) = f(x)^T \beta = \beta_0 + x_{1i}^T \beta_1 + x_{2i}^T \beta_2 \tag{3-4}$$

该模型的信息矩阵为

$$M(\beta; x_i) = \sum_i^n w_i v_i f(x) f(x)^T = \begin{pmatrix} \sum_{i=1}^n w_i v_i & \sum_{i=1}^n w_i v_i x_{1i} & \sum_{i=1}^n w_i v_i x_{2i} \\ \sum_{i=1}^n w_i v_i x_{1i} & \sum_{i=1}^n w_i v_i x_{1i}^2 & \sum_{i=1}^n w_i v_i x_{1i} x_{2i} \\ \sum_{i=1}^n w_i v_i x_{2i} & \sum_{i=1}^n w_i v_i x_{1i} x_{2i} & \sum_{i=1}^n w_i v_i x_{2i}^2 \end{pmatrix} \tag{3-5}$$

记 $v_i = \dfrac{e^{\beta_0+\beta_1 x_{1i}+\beta_2 x_{2i}}}{(1+e^{\beta_0+\beta_1 x_{1i}+\beta_2 x_{2i}})^2}$，故两变量模型信息矩阵的逆矩阵为

$$M^{-1}(\beta; x_i) = \frac{1}{\sum_{i=1}^n w_i v_i h_{11} - \sum_{i=1}^n w_i v_i x_{1i} h_{12} + \sum_{i=1}^n w_i v_i x_{2i} h_{13}} \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \tag{3-6}$$

其中

$$h_{11} = \sum_{i=1}^n w_i v_i x_{1i}^2 \cdot \sum_{i=1}^n w_i v_i x_{2i}^2 - \left(\sum_{i=1}^n w_i v_i x_{1i} x_{2i} \right)^2, h_{12} = \sum_{i=1}^n w_i v_i x_{2i} \cdot \sum_{i=1}^n w_i v_i x_{1i} x_{2i} - \sum_{i=1}^n w_i v_i x_{1i} \cdot \sum_{i=1}^n w_i v_i x_{2i}^2$$

$$h_{13} = \sum_{i=1}^n w_i v_i x_{1i} \cdot \sum_{i=1}^n w_i v_i x_{1i} x_{2i} - \sum_{i=1}^n w_i v_i x_{2i} \cdot \sum_{i=1}^n w_i v_i x_{1i}^2, h_{21} = \sum_{i=1}^n w_i v_i x_{1i} x_{2i} \cdot \sum_{i=1}^n w_i v_i x_{2i} - \sum_{i=1}^n w_i v_i x_{1i} \cdot \sum_{i=1}^n w_i v_i x_{2i}^2$$

$$h_{22} = \sum_{i=1}^n w_i v_i \cdot \sum_{i=1}^n w_i v_i x_{2i}^2 - \left(\sum_{i=1}^n w_i v_i x_{2i} \right)^2, h_{23} = \sum_{i=1}^n w_i v_i x_{1i} \cdot \sum_{i=1}^n w_i v_i x_{2i} - \sum_{i=1}^n w_i v_i \cdot \sum_{i=1}^n w_i v_i x_{1i} x_{2i}$$

$$h_{31} = \sum_{i=1}^n w_i v_i x_{1i} \cdot \sum_{i=1}^n w_i v_i x_{1i} x_{2i} - \sum_{i=1}^n w_i v_i x_{1i} x_{2i} \cdot \sum_{i=1}^n w_i v_i x_{2i}, h_{32} = \sum_{i=1}^n w_i v_i x_{1i} \cdot \sum_{i=1}^n w_i v_i x_{2i} - \sum_{i=1}^n w_i v_i \cdot \sum_{i=1}^n w_i v_i x_{1i} x_{2i}$$

$$h_{13} = \sum_{i=1}^n w_i v_i \cdot \sum_{i=1}^n w_i v_i x_{1i}^2 - \left(\sum_{i=1}^n w_i v_i x_{1i} \right)^2 。$$

上述我们得到了固定截距 Logistic 模型的信息矩阵且假定该矩阵是可逆的，由此便可以构造出相应的最优设计函数。

3.2 固定截距模型的 R-最优设计

已知前文 2.2 节阐述了 Logistic 回归模型 R-最优设计的准则函数及一般等价性定理，根据上节介绍的固定截距单变量 Logistic 回归模型的信息矩阵，该模型具体的准则函数为

$$\psi(M(\xi))=\prod_{j=1}^{p}e_j^T M^{-1}(\xi)e_j=\prod_{j=1}^{p}e_j^T\left\{\frac{\begin{pmatrix}\sum_{i=1}^{n}w_i x_{1i}v_i x_{1i} & -\sum_{i=1}^{n}w_i v_i x_{1i}\\ -\sum_{i=1}^{n}w_i x_{1i}v_i & \sum_{i=1}^{n}w_i v_i\end{pmatrix}}{\sum_{i=1}^{n}w_i v_i\cdot\sum_{i=1}^{n}w_i x_{1i}v_i x_{1i}-\left(\sum_{i=1}^{n}w_i v_i x_{1i}\right)^2}\right\}e_j \tag{3-7}$$

$$=\frac{\sum_{i=1}^{n}w_i x_{1i}v_i x_{1i}\cdot\sum_{i=1}^{n}w_i v_i}{\left(\sum_{i=1}^{n}w_i v_i\cdot\sum_{i=1}^{n}w_i x_{1i}v_i x_{1i}-\left(\sum_{i=1}^{n}w_i v_i x_{1i}\right)^2\right)^2}$$

我们称使得信息矩阵 $M(\xi)$ 非退化的设计 ξ_R^* 为模型（3-1）的 R-最优设计则需要满足以下要求：

$$\psi(M(\xi_R^*))=\min_{\xi\in\Xi}\prod_{j=1}^{p}(M^{-1}(\xi))_{jj}$$

$$=\min_{\xi\in\Xi}\frac{\sum_{i=1}^{n}w_i x_{1i}\frac{e^{\beta_0+\beta_1 x_{1i}}}{(1+e^{\beta_0+\beta_1 x_{1i}})^2}x_{1i}\cdot\sum_{i=1}^{n}w_i\frac{e^{\beta_0+\beta_1 x_{1i}}}{(1+e^{\beta_0+\beta_1 x_{1i}})^2}}{\left(\sum_{i=1}^{n}w_i\frac{e^{\beta_0+\beta_1 x_{1i}}}{(1+e^{\beta_0+\beta_1 x_{1i}})^2}\cdot\sum_{i=1}^{n}w_i x_{1i}\frac{e^{\beta_0+\beta_1 x_{1i}}}{(1+e^{\beta_0+\beta_1 x_{1i}})^2}x_{1i}-\left(\sum_{i=1}^{n}w_i\frac{e^{\beta_0+\beta_1 x_{1i}}}{(1+e^{\beta_0+\beta_1 x_{1i}})^2}x_{1i}\right)^2\right)^2} \tag{3-8}$$

$$=\min_{\xi\in\Xi}\frac{\sum_{i=1}^{n}w_i x_{1i}v_i x_{1i}\cdot\sum_{i=1}^{n}w_i v_i}{\left(\sum_{i=1}^{n}w_i v_i\cdot\sum_{i=1}^{n}w_i x_{1i}v_i x_{1i}-\left(\sum_{i=1}^{n}w_i v_i x_{1i}\right)^2\right)^2}$$

同样两变量条件下固定截距的 logistic 回归模型具体的准则函数为

$$\psi(M(\xi))=\prod_{j=1}^{p}e_j^T M^{-1}(\xi)e_j=\frac{h_{11}h_{22}h_{33}}{\left(\sum_{i=1}^{n}w_i v_i h_{11}-\sum_{i=1}^{n}w_i v_i x_{1i}h_{12}+\sum_{i=1}^{n}w_i v_i x_{2i}h_{13}\right)^3}$$

该准则函数满足 R-最优设计的条件为

$$\psi(M(\xi_R^*))=\min_{\xi\in\Xi}\psi(M(\xi))=\min_{\xi\in\Xi}\frac{h_{11}h_{22}h_{33}}{\left(\sum_{i=1}^{n}w_iv_ih_{11}-\sum_{i=1}^{n}w_iv_ix_{1i}h_{12}+\sum_{i=1}^{n}w_iv_ix_{2i}h_{13}\right)^3} \quad (3\text{-}9)$$

在节 2.3 中已经给出了 Logistic 回归模型的 R-最优设计等价的敏感性函数，模型 R-最优设计的求解可以等价的转化为求解使得以下敏感性函数 $\phi(x,\xi)$ 在设计域上的值最接近 P 的设计，例如单因素情况，在模型（3-1）下：

$$\phi(x,\xi)=tr\left\{v(x)f(x)^TM^{-1}(\xi)\sum_{j=1}^{p}\frac{e_j^Te_j}{e_j^TM^{-1}(\xi)e_j}M^{-1}(\xi)f(x)\right\} \quad (3\text{-}10)$$

其中 $\sum_{j=1}^{p}\frac{1}{e_j^TM^{-1}(\xi^*)e_j}=\frac{c_{01}}{\sum_{i=1}^{n}w_ix_{1i}v_ix_{1i}}+\frac{c_{02}}{\sum_{i=1}^{n}w_iv_i}$，$c_{01}=\sum_{i=1}^{n}w_iv_i\cdot\sum_{i=1}^{n}x_{1i}w_iv_ix_{1i}-\left(\sum_{i=1}^{n}w_iv_ix_{1i}\right)^2$

使得上面函数的上界最接近 2 时的设计即为 R-最优设计。

同样两因素情况，在模型（3-4）下，$\phi(x,\xi)$ 函数为：

$$\phi(x,\xi)=tr\left\{v(x)f(x)^TM^{-1}(\xi)\sum_{j=1}^{p}\frac{e_j^Te_j}{e_j^TM^{-1}(\xi)e_j}M^{-1}(\xi)f(x)\right\} \quad (3\text{-}11)$$

其中

$$\sum_{j=1}^{p}\frac{1}{e_j^TM^{-1}(\xi^*)e_j}=\frac{c_{20}}{h_{11}}+\frac{c_{20}}{h_{22}}+\frac{c_{20}}{h_{33}}，\ c_{20}=\sum_{i=1}^{n}w_iv_ih_{11}-\sum_{i=1}^{n}w_iv_ix_{1i}h_{12}+\sum_{i=1}^{n}w_iv_ix_{2i}h_{13}$$

使得上面函数的上界最接近 3 时的设计即为 R-最优设计。

由于对于不同设计定义的参数范围不完全相同，为了区分出不同设计效果的优劣性，我们定义一个判断的指标，即 R-效率。

在设计域 χ 中，设计 ξ 在模型（3-1）中的 R-效率为

$$Eff_R(\xi)=\frac{\psi_R(M(\xi_R^*))}{\psi_R(M(\xi))} \quad (3\text{-}12)$$

其中 ξ_R^* 为该模型的 R-最优设计，显然在设计域中 $Eff_R(\xi)$ 的取值为 0 到 1。需要注意 $\psi\left(M(\xi_R^*)\right)$ 的取值与未知参数的选取有关，故模型（3-1）的准则函数不存在绝对意义上的单调性。

3.3 R-最优准则的迭代算法

求解本文 Logistic 回归模型 R-最优的迭代算法具体如下：

- 首先给定一个足够小的阈值 T，T>0, 任意取一个非退化的初始设计 ξ_0，设计点个数为 n，记迭代的次数 k=1
- 然后计算该设计的信息矩阵 $M(\beta;\xi_0)$，求出函数 $\phi(x,\xi)$，若

$$\left|\sup_{x\in\chi}\phi(x,\xi_0)-p\right|\le T,$$

则取 R-最优设计 $\xi^*=\xi_0$；否则通过求解得到满足该设计最大的支撑点 x_k^*

$$x_k^*=\arg\sup_{x\in\aleph}\phi(x,\xi_k)$$

- 接着把点 x_k^* 处权重为 1 的单点测度记为 $\xi_{x_k^*}^-$，求解在 $\xi_{x_k^*}^-$ 处的步长

$$\alpha_k=\arg\min_{\alpha}\psi_R(1-\alpha)\xi_k+\alpha\xi_{x_k^*}$$

- 之后根据第三步得到的 $\xi_{x_k^*}$ 和 α_k，我们可以更新设计进行迭代：

$$\xi_{k+1}=(1-\alpha_k)\xi_k+\alpha_k\xi_{x_k^*}$$

- 此次迭代结束，继续下一次迭代，记 k=k+1，重复第二步，最终求解出 R-最优结果。为了使算法简化，本文采用的步长为

$$\alpha_k^*=\frac{1}{k}$$

3.4 数值模拟

Linda 等[14]给出了 Logistic 固定截距两因素模型的 D-最优设计在三个设计点和四个设计点下的局部最优设计解，本节主要举例求解出不同变量个数 Logistic 固定截距模型在假定参数范围-4 到 2 的 R-最优设计。

例 3.4.1 对于单变量的固定截距 Logistic 回归模型求解局部 R-最优设计，假定设计域 $\chi=[0,1]$，设计点 $x_i\in\chi$，取初始设计 $\xi_{1,0}=\begin{pmatrix}0 & 1\\ 0.5 & 0.5\end{pmatrix}$，阈值 T 选取 0.005，参数 $\beta=(1,\beta_1)^T=(1,1)^T$，则函数为 $y_i=1+x_i$。

当迭代次数超过 5000 时跳出循环，取 y=$p-\phi(x,\xi)$，最终我们可以得到该设定下单因素固定截距模型的 R-最优设计为

$$\xi_R^*=\begin{pmatrix} 0 & 1 \\ 0.6223 & 0.3777 \end{pmatrix}$$

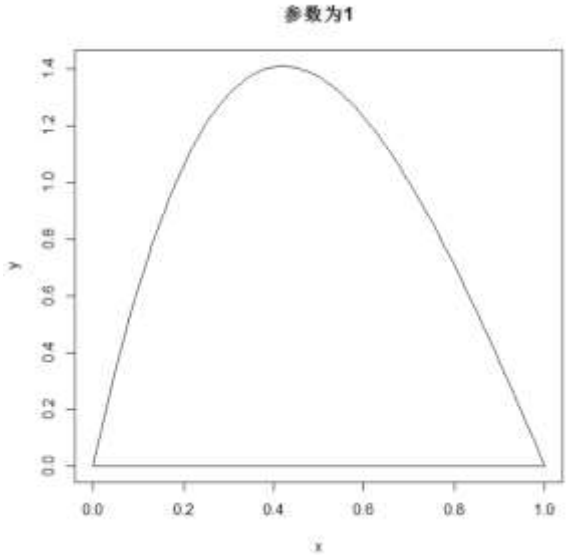


图 3-1 β_1取 1 时 y 与 x 的关系

当设计为ξ_R^*时，上图为 R-最优设计迭代中 y 与 x 的关系图，显然 $p-\phi(x,\xi)$ 在 $x=0$ 和 $x=1$ 处取得值 0，即在 $x=0$ 和 $x=1$ 处为 $p-\phi(x,\xi)$ 的零值点。

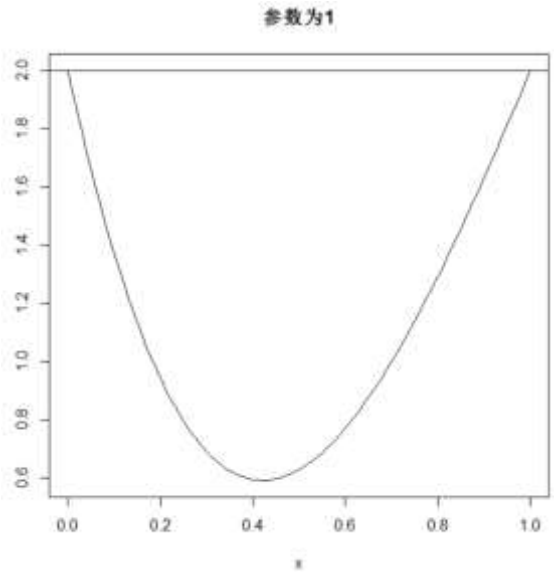


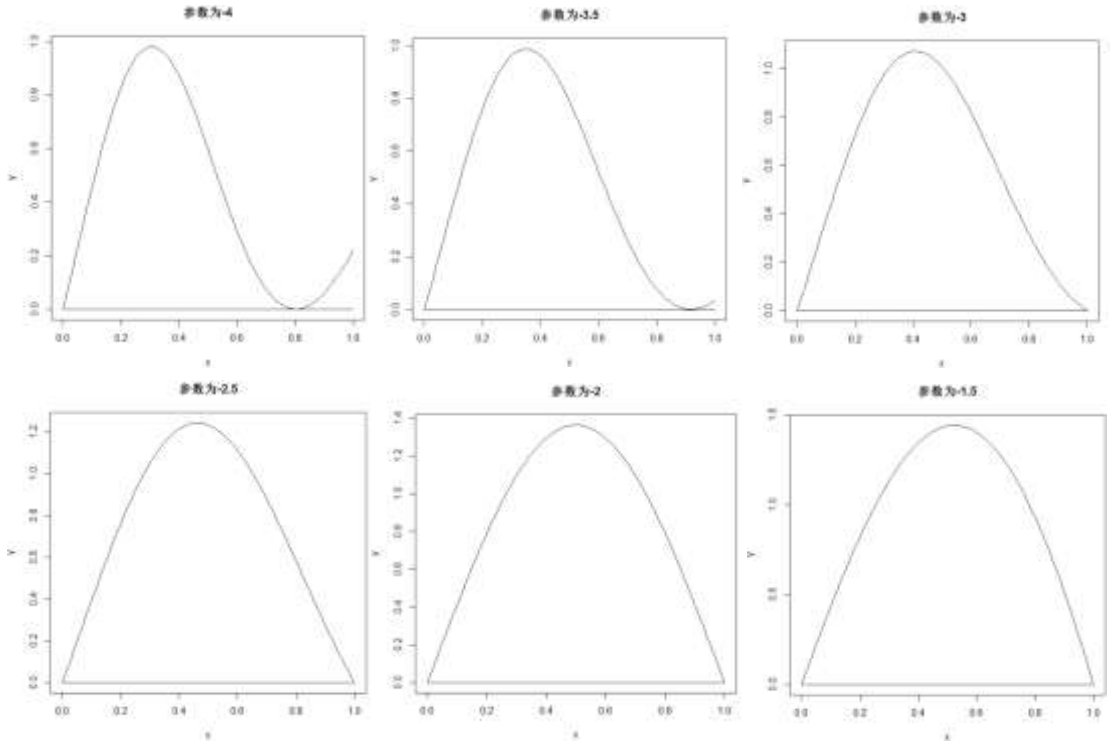
图 3-2 β_1取 1 时 $\phi(x,\xi_R^*)$ 与 x 的关系

当 β_1=1时，函数 $\phi(x,\xi_R^*)$ 为最大值 2，故认为设计 ξ_R^*为该模型的 R-最优设计。

表 3-1 β_1取不同值时单因素模型的 R-最优设计点

序号	β_1	$X_{(m)}$	$P\left(X_{(m)}\right)$	$(2\text{-}\phi(\xi))\|x=1$
1	-4	0,0.8	0.6113 ,0.3887	0.219021
2	-3.5	0,0.915	0.6119,0.3881	0.028468
3	-3	0,1	0.6223 0.3777	0.000474
4	-2.5	0,1	0.6465,0.3535	0.000507
5	-2	0,1	0.6667,0.3333	-0.000300
6	-1.5	0,1	0.6799,0.3201	0.000928
7	-1	0,1	0.6847,0.3153	-0.000322
8	-0.5	0,1	0.6799,0.3201	0.000928
9	0.5	0,1	0.6465,0.3535	0.000507
10	1	0,1	0.6223,0.3777	0.000474
11	1.5	0,1	0.5969,0.4031	0.000653
12	2	0,1	0.5731,0.4269	-0.000129

当初始参数值取值为-4 到 2 时，假定函数$\phi(x,\xi)$与 p 值的差为 y，其与 x 的关系如下所示：



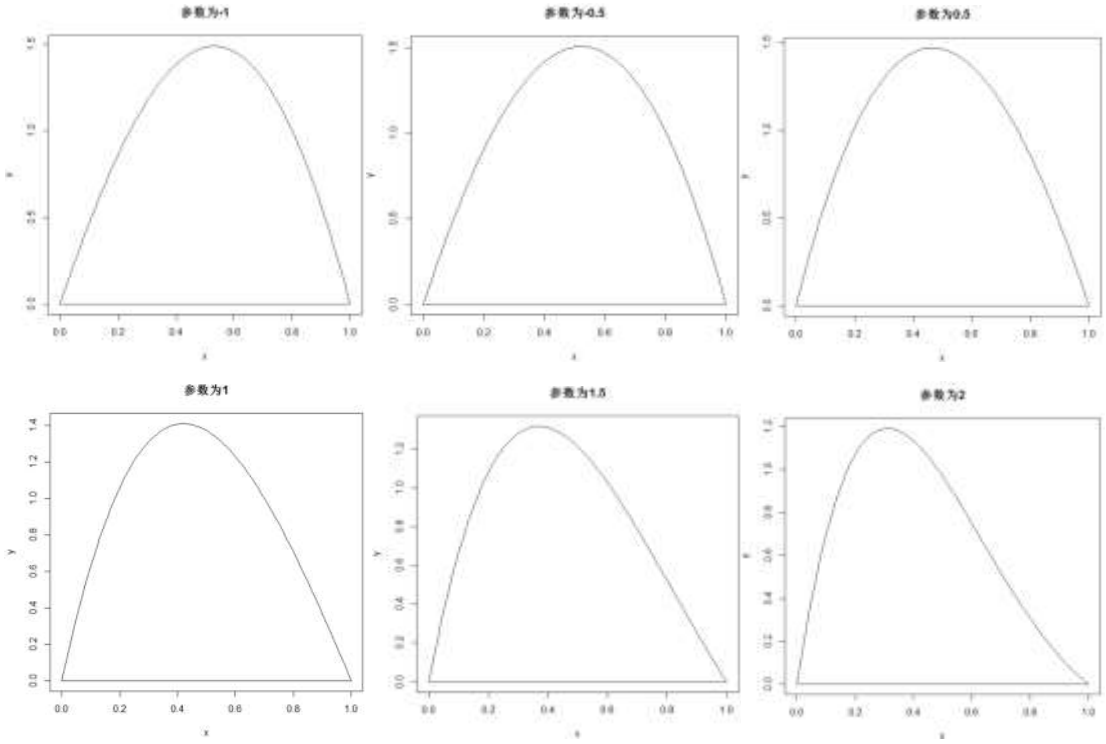


图 3-3 β_1取-4 至 2 时 y 与 x 的关系

故当β_1=-4，β_1=-3.5，β_1=-3，β_1=-2.5 β_1=-2，β_1=-1.5，β_1=-1，β_1=-0.5，β_1=0.5，β_1=1，β_1=1.5 β_1=2 时，在初始设计下找到了模型的 R-最优设计，即表 3-1 中的设计为ξ_R^*设计。

在初始设计$\xi_{1,0}=\begin{pmatrix} 0 & 1 \\ 0.5 & 0.5 \end{pmatrix}$下，参数$\beta_1$不同取值的 R-最优设计效率为

表 3-2 不同参数取值固定截距单因素模型的 R 效率

序号	β_1	$Eff_R(\xi)$
1	-4	0.8690
2	-3.5	0.9364
3	-3	0.9181
4	-2.5	0.8801
5	-2	0.8437
6	-1.5	0.8177
7	-1	0.8083

8	-0.5	0.8177
9	0.5	0.8801
10	1	0.9181
11	1.5	0.9504
12	2	0.9731

例 3.4.2 对于两变量的固定截距 Logistic 回归模型求解局部 R-最优设计，我们假定设计域 $\chi=[0,2]\times[0,2]$，设计点 $x_{1i},x_{2i}\in\chi$，取初始九点设计

$$\xi_{2,0}=\begin{pmatrix}(0,0) & (1,0) & (2,0) & (0,1) & (1,1) & (2,1) & (0,2) & (1,2) & (2,2)\\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9}\end{pmatrix}$$

阈值选取 $T=0.005$，参数 $\beta=(1,\beta_1,\beta_2)^T=(1,1,1)^T$，当迭代次数超过 5000 时跳出循环。该初始设计 $\xi_{2,0}$ 在两因素固定截距模型的 R-最优设计为

$$\xi_R^*=\begin{pmatrix}(0,0) & (2,0) & (0,2)\\ 0.4234 & 0.2883 & 0.2883\end{pmatrix}$$

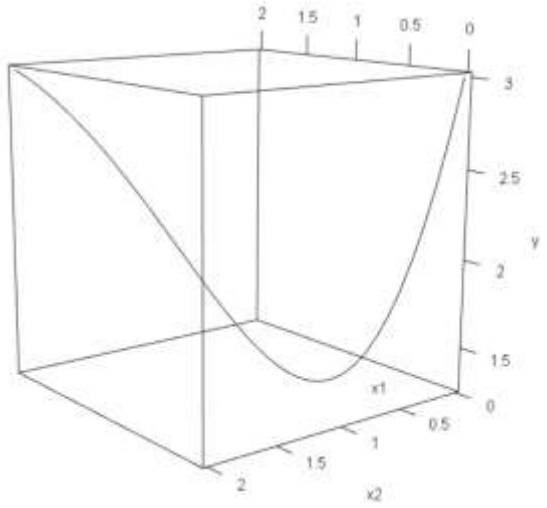


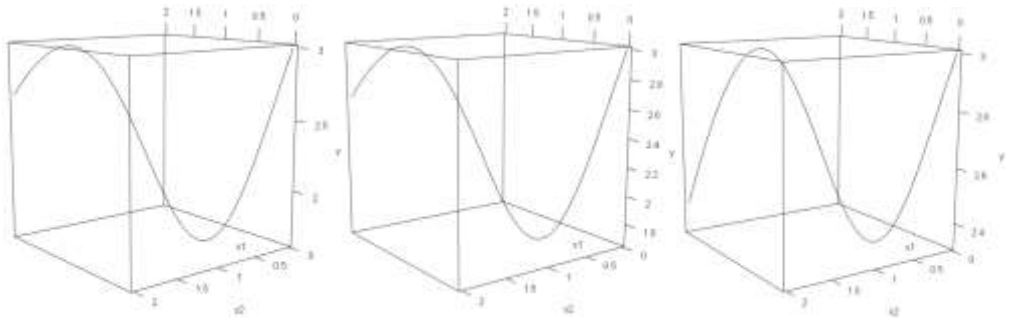
图 3-4 β_1，β_2 取(1,1)时 $\phi(x,\xi)$ 与 x 的关系

在设计域 $[0,2]^2$ 上，显然 $\phi(x,\xi)$ 函数的最大值在 ξ_R^* 处取得且值为 3，因此设计 ξ_R^* 为两因素模型在参数初值 $(1,1,1)^T$ 模型的 R-最优设计。

表 3-3 β 取不同值时两因素模型的 R-最优设计点

序号	β_1	β_2	$X_{(m)}$	$P(X_{(m)})$
1	-2	-2	(0,0),(1.585,0)(0,1.582)	0.4718 ,0.2628,0.2634
2	-2	-1	(0,0),(0,2),(1.577,0)	0.5042,0.2292,0.2646
3	-2	1	(0,0),(2,0), (1.406,0),(0.5855,2)	0.379,0.1058, 0.2738,0.237
4	-2	2	(0,0),(2,2)(1.703,2) (1.575,0)(0,0.2892)(2,0.417 1)	0.4238,0.061,0.1092, 0.3644,0.007,0.0282
5	-1	-2	(0,0),(2,0),(0,1.577)	0.5042,0.2292,0.2646
6	-1	-1	(0,0),(2,2),(0,2), (1.289,2),(2,1.291)	0.5142,0.2148,0.2148, 0.0262,0.0272
7	-1	1	(0,0),(2,0),(2,2),(0.7161,2)	0.4490,0.2932,0.1590,0.0966
8	-1	2	(0,0),(2,0),(2,1.653)	0.4622,0.2916,0.2442
9	1	-2	(0,0),(2,2),(0,1.4057),(2,0.5 855)	0.379,0.1058,0.2738,0.2370
10	1	-1	(0,0),(0,2),(2,2),(2,0.7194)	0.449,0.2932,0.159,0.0966
11	1	1	(0,0),(2,0),(0,2)	0.4234,0.2883,0.2883
12	1	2	(0,0),(2,0),(0,1.193)	0.4138,0.2874,0.2974

当初始参数值取值不同时，$\phi(x,\xi)$ 与 x 的关系如下所示：



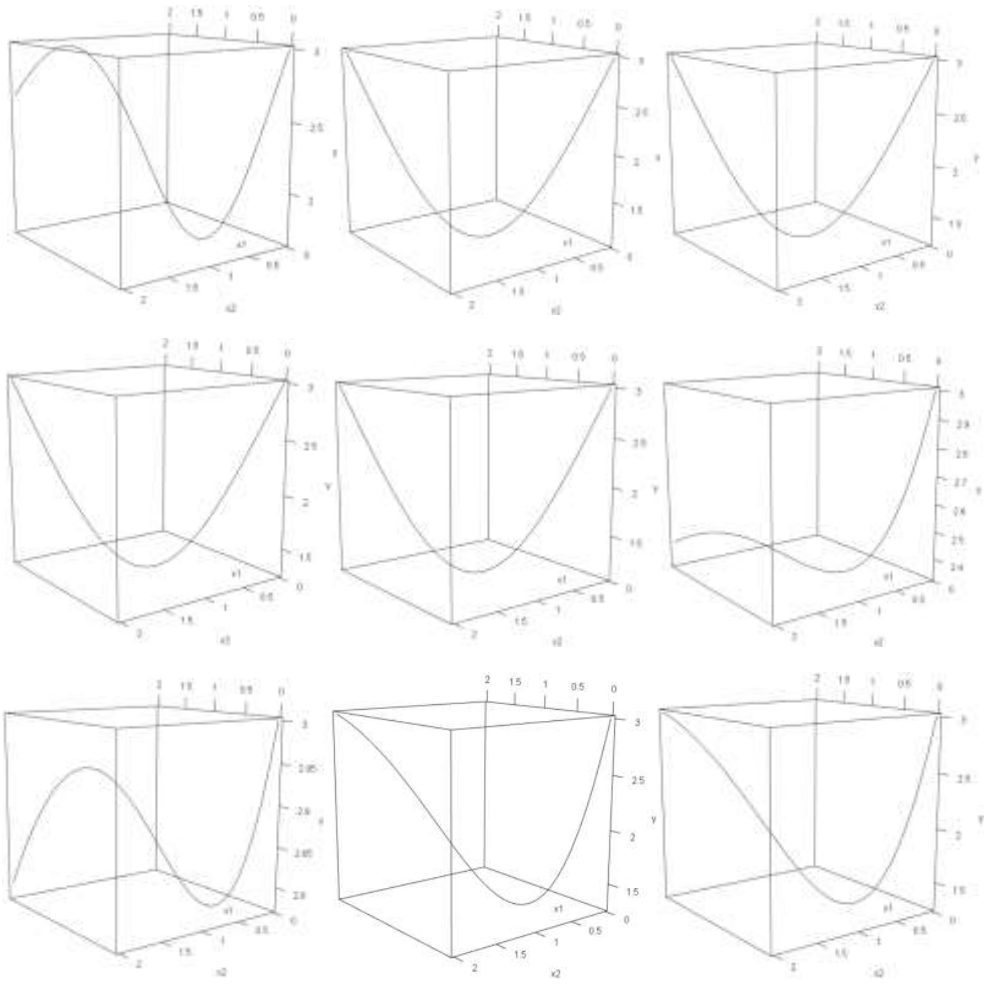


图 3-5 β 取-2 至 2 时 $\phi(x,\xi)$ 与 x 的关系

在初始设计 $\xi_{2,0}$ 下，参数 β 不同取值的 R-最优设计效率分别为

表 3-4 不同参数取值固定截距两因素模型的 R-效率

序号	β_1	β_2	$Eff_R(\xi)$
1	-2	-2	0.1162
2	-2	-1	0.1661
3	-2	1	0.6012
4	-2	2	0.2977
5	-1	-2	0.1662
6	-1	-1	0.8985
7	-1	1	0.3018

8	-1	2	0.1854
9	1	-2	0.3772
10	1	-1	0.3018
11	1	1	0.1722
12	1	2	0.1377

第四章 随机截距的 Logistic 回归模型

4.1 随机截距模型及信息矩阵

4.1.1 单变量模型

在单因素情况下，Logistic 随机截距回归模型可以写为

$$\log itP(y_{ij}=1|x_{ij},b_j)=\log it(\pi_{ij})=\ln(\frac{\pi_{ij}}{1-\pi_{ij}})=x_{ij}^T\beta+b_j \qquad (4\text{-}1)$$

上式 y_{ij} 表示随机截距模型第 j 层（$j=1,2,\cdots,m$）第 i 次观测（$i=1,2,\cdots,n$）的观测值。其中 $x_{ij}{}^T=(1,x_{1ij})$，$\beta^T=(\beta_0,\beta_1)$，$\pi_{ij}=P(y_{ij}=1|x_{ij})=\frac{e^{x_{ij}^T\beta+b_j}}{1+e^{x_{ij}^T\beta+b_j}}$，$b_j$ 代表不同层的随机部分且相互独立，且 b_j 服从均值为 0 方差为 σ_b^2 的正态分布。

记 $f_{ij}=(1,x_{ij})^T$，$F_j=(f_{1j},\ldots,f_{nj})^T$，$Y_j^*=(y_{1j}^*,y_{2j}^*,\ldots,y_{nj}^*)^T$，$\pi_j=(\pi_{1j},\ldots,\pi_{nj})^T$，上述模型第 j 层的全体观测值对应的线性回归模型可表示为

$$Y_j^* \triangleq \log it(\pi_j)=F_j\beta+1_n b_j,$$

所有线性回归模型的全体可表示为

$$Y^* \triangleq \begin{pmatrix}F_1\\ \vdots\\ F_m\end{pmatrix}\beta+\begin{pmatrix}1_n & & 0\\ & \ddots & \\ 0 & & 1_n\end{pmatrix}\begin{pmatrix}b_1\\ \vdots\\ b_m\end{pmatrix}=F\beta+I_{nm}b. \qquad (4\text{-}2)$$

根据 2.2.1 节构造的准信息矩阵可知单因素情况下为

$$M(\beta;F_j)=Cov(\hat{\beta})_j^{-1}=F_j^TV_j^{-1}F_j. \qquad (4\text{-}3)$$

其中 $V_j=\begin{pmatrix}(\pi_{1j}(1-\pi_{1j}))^{-1}+\sigma_b{}^2 & \sigma_b{}^2 & \cdots & \sigma_b{}^2\\ \sigma_b{}^2 & (\pi_{2j}(1-\pi_{2j}))^{-1}+\sigma_b{}^2 & \cdots & \sigma_b{}^2\\ \cdots & \cdots & \ddots & \vdots\\ \sigma_b{}^2 & \sigma_b{}^2 & \cdots & (\pi_{nj}(1-\pi_{nj}))^{-1}+\sigma_b{}^2\end{pmatrix}$，$F_j=\begin{pmatrix}1 & x_{1j}\\ \vdots & \\ 1 & x_{nj}\end{pmatrix}$。

接着定义每层近似设计 ξ 的概率分布，假设设计 $\xi\in\Xi$，有

$$\xi=\begin{Bmatrix}X_{(1j)} & \cdots & X_{(nj)}\\ w_1 & \cdots & w_n\end{Bmatrix},X_{(ij)}=(x_{ij1},\cdots,x_{ijp}),0<w_i<1,\sum_{i=1}^{n}w_i=1,$$

故整体模型设计构造表示如下：

表 4-1 Logistic 随机截距模型的具体参数定义

群组	权重	观测点	随机变量	每层权重
1	w	$x_{11},\cdots,x_{n1}$	$b_1,\cdots,b_1$	w_1 ⋮ w_n
⋮	⋮	⋮	⋮	⋮
m	w	$x_{1m},\cdots,x_{nm}$	$b_m,\cdots,b_m$	w_1 ⋮ w_n

4.2 简单随机截距模型的 R-最优设计

取 n=2，在设计 ξ 中单变量随机截距模型的信息准则矩阵为

$$M(\beta;\xi)\approx\frac{1}{a_1a_2-b^2}\begin{pmatrix}a_1+a_2-2b & (a_2-b)w_{1j}x_{1j}+(a_1-b)w_{2j}x_{2j}\\ (a_2-b)w_{1j}x_{1j}+(a_1-b)w_{2j}x_{2j} & a_2w_{1j}x_{1j}^2+a_1w_{2j}x_{2j}^2-2bw_{ij}x_{1j}x_{2j}\end{pmatrix} \quad (4\text{-}4)$$

其中 $a_i=(\pi_{ij}(1-\pi_{ij}))^{-1}+{\sigma_b}^2$，$b={\sigma_b}^2$，模型 R-最优设计具体的准则函数可以写为

$$\psi(\xi)=\prod_{k=1}^{p}e_k^TM^{-1}(\xi)e_k=\frac{(a_1+a_2-2b)}{\left(a_1a_2-b^2\right)^2}(a_2w_{1j}x_{1j}^2+a_1w_{2j}x_{2j}^2-2bw_{ij}x_{1j}x_{2j}) \quad (4\text{-}5)$$

同样我们称使得信息矩阵 $M(\xi)$ 非退化的设计 ξ_R^* 为模型（4-2）的 R-最优设计则需要满足以下要求：

$$\psi(\xi_R^*)=\min_{\xi\in\Xi}\prod_{k=1}^{p}(M^{-1}(\xi))_{kk}$$

由于计算求解有限，本文随机截距仅考虑最基础情况，即 j=1。由于只有一层的 Logistic 随机截距模型等价于在固定截距情况上增加了一个定值变量，仍属于固定截距范畴。根据前文固定截距情况下 R-最优设计的一般等价性定理，假设

$$o_i=\frac{e^{\beta_0+\beta_1x_{ij}+b_j}}{(1+e^{\beta_0+\beta_1x_{ij}+b_j})^2},$$

$$c=\sum_{i=1}^{n}w_{ij}\left(\frac{e^{\beta_0+\beta_1x_{ij}+b_j}}{(1+e^{\beta_0+\beta_1x_{ij}+b_j})^2}\right)\cdot\sum_{i=1}^{n}x_{ij}w_{ij}\left(\frac{e^{\beta_0+\beta_1x_{ij}+b_j}}{(1+e^{\beta_0+\beta_1x_{ij}+b_j})^2}\right)x_{ij}-\left(\sum_{i=1}^{n}w_{ij}\left(\frac{e^{\beta_0+\beta_1x_{ij}+b_j}}{(1+e^{\beta_0+\beta_1x_{ij}+b_j})^2}\right)x_{ij}\right)^2,$$

可以得到单层单变量随机截距模型的 $\phi(x,\xi)$ 函数

$$\begin{aligned}&\phi(x,\xi)=tr\left\{o_i f(x)^T M^{-1}(\xi)\sum_{i=1}^{p}\frac{e_i^T e_i}{e_i^T M^{-1}(\xi)e_i}M^{-1}(\xi)f(x)\right\},\\&\sum_{k=1}^{p}\frac{1}{e_i^T M^{-1}(\xi)e_i}=\frac{c}{\sum_{i=1}^{n}w_{ij}x_{1i}o_i x_{ij}}+\frac{c}{\sum_{i=1}^{n}w_{ij}o_i}.\end{aligned}\qquad(4\text{-}6)$$

4.3 数值模拟

本节主要举例求解出单层单变量 Logistic 随机截距模型在假定参数范围-2 到 2.5 的 R-最优设计。

例 4.4.1 对于单变量的随机截距 Logistic 回归模型求解局部 R-最优设计，假定设计域 $\chi=[0,1]$，设计点 $x_{1i}\in\chi$，从简单角度考虑取初始设计 $\xi_{3,0}=\begin{pmatrix}0 & 1\\0.5 & 0.5\end{pmatrix}$，阈值选取 $T=0.005$，参数 $\beta=(1,\beta_1)^T=(1,1)^T$，层数 j=1，随机系数 b_1 服从 $N(0,1)$ 分布，则模型函数为 $y_i=1+x_i+b_1$。当迭代次数超过 5000 时跳出循环，同样取 $y=p-\phi(x,\xi)$，最终我们可以得到以上设定下单因素随机截距模型的 R-最优设计为

$$\xi_R^*=\begin{pmatrix}0 & 1\\0.6431 & 0.3569\end{pmatrix}$$

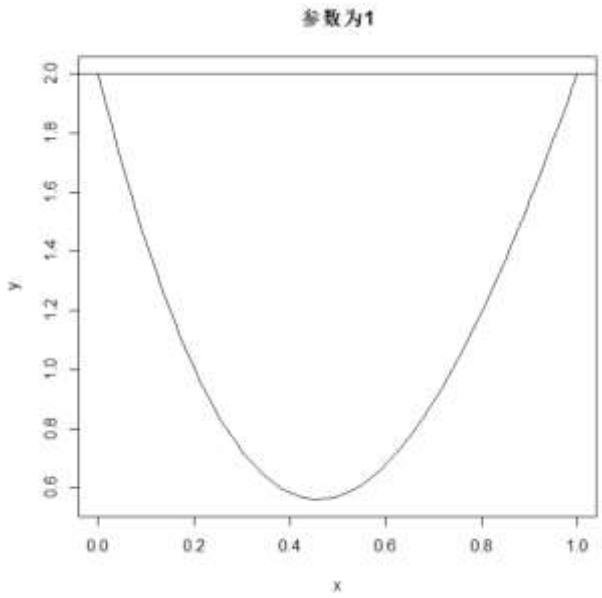


图 4-1 β 取-1 时 $\phi(x,\xi)$ 与 x 的关系

不难看出，在[0,1]上 $\phi(x,\xi)$ 函数的最大值为 2 且在设计 ξ_R^* 处取得，因此设计 ξ_R^* 为单因素单层随机系数模型在参数初值取 $(1,1)^T$，随机系数初值服从 $N(0,1)$ 分布的 R-最优设计。

表 4-2 β_1 取不同值时随机截距单因素模型的 R-最优设计点

序号	β_1	$X_{(m)}$	$P(X_{(m)})$	$\phi(x,\xi)\|x=1$
1	-2	0,1	0.6843,0.3157	2.0003
2	-1.5	0,1	0.7077,0.2923	1.9999
3	-1	0,1	0.6789,0.3211	1.9996
4	-0.5	0,1	0.6943,0.3057	2.0002
5	0.5	0,1	0.6485,0.3515	1.9999
6	1	0,1	0.6431,0.3569	1.9998
7	1.5	0,1	0.6187,0.3813	1.9997
8	2	0,1	0.6131,0.3869	1.9999
9	2.5	0,1	0.5757,0.4243	2.0001

当初始参数值取值为-2 到 2.5 时，$\phi(x,\xi)$ 与 x 的关系如下所示：

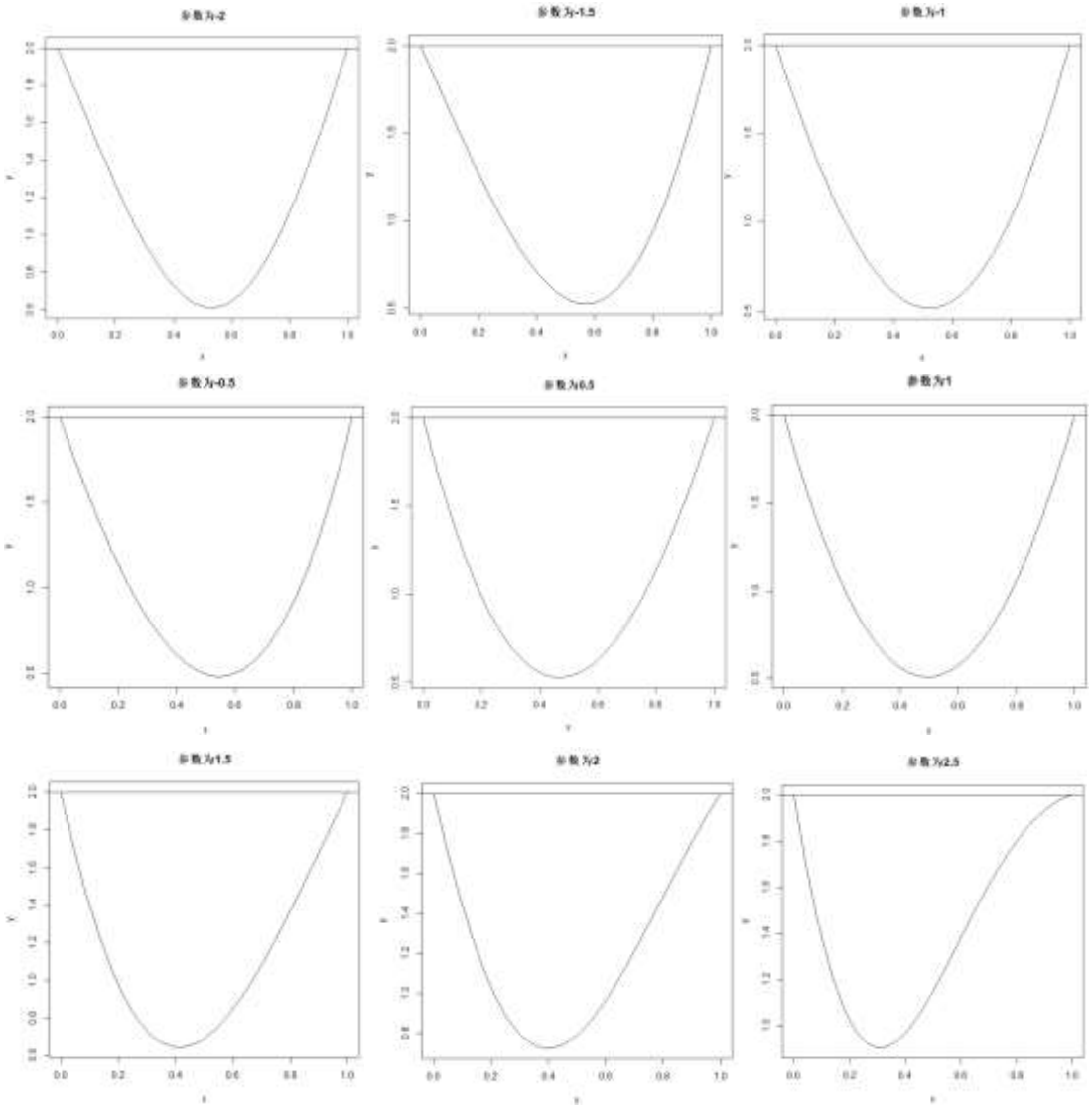


图 4-2 β_1取-2 至 2.5 时 $\tilde{\phi}(x,\xi)$ 与 x 的关系

在参数取值为 $\beta=(1,\beta_1)^T=(1,-2)^T$ 和 $\beta=(1,\beta_1)^T=(1,1)^T$ 时，单因素 Logistic 随机截距回归模型最优设计的 R-效率分别为

$$Eff_R(\xi)_{-2}=\frac{194.5262}{240.4151}\times 100\%\approx 80.91\%.$$

$$Eff_R(\xi)_1=\frac{183.3399}{202.4441}\times 100\%\approx 90.56\%.$$

第五章 结 论

本文主要讨论了 Logistic 回归模型在固定截距和随机截距条件下的 R-最优设计。已知线性模型 R-最优算法的基础上，通过一般等价性定理推导出 R-最优设计的敏感性函数，由于敏感性函数中的信息矩阵包含待估参数，故本文尝试先假定未知参数的初值范围，再构造出对应迭代求解的算法，最终获得相关模型的局部 R-最优设计。

在固定截距的 Logistic 回归模型中，首先考虑单因素模型的 R-最优准则，假设参数β_1的取值为-4 到 2，根据构造的方向导数以及敏感性函数，通过迭代算法求解单因素固定截距 logistic 回归模型的两个设计点的 R-最优设计，并且利用函数$\phi(x,\xi)$的图像来验证设计的最优性与准确性，最后计算出 R 效率。其次考虑双因素模型，在单因素模型的基础上维度增加，假设参数β_1的取值为-2 到 1，β_2的取值为-2 到 2，同样利用$\phi(x,\xi)$函数与迭代算法求解出两个设计点的 R-最优设计，最后用敏感性函数图像来加以验证，计算出模型的 R-效率。

在随机截距的 Logistic 回归模型中，从最简单的情况入手，只考虑单因素模型的 R-最优设计，首先选取参数β_1的初值为-2 到 2.5，随机截距假设服从$N(0,1)$分布，在单层初始设计下，通过迭代算法求解单因素随机截距 Logistic 回归模型的 R-最优设计，画出$\phi(x,\xi)$函数图像进行验证。

结果表明，本文中固定截距与随机截距 Logistic 模型的 R-最优设计都是存在的，并且计算得到的单因素模型以及双因素模型 R-效率均体现了 R-最优设计相比初始设计而言模型参数效果更好，同时简单随机截距模型在参数值取 1 时的 R-效率为 90.56%。

由于个人理论知识的局限以及程序实现的时效性，在模型构造以及数值列举方面都是从较为简单的层面考虑，涉及复杂高维模型的部分还没有详细说明。在今后的学习中，针对复杂混合 Logistic 回归模型的局部 R-最优设计可以进一步研究讨论。

参考文献

[1] Fedorov V V. Theory of optimal experiments[M]. Academic Press, 1972, 59(3):345-350.

[2] Fedorov V V. Convex design theory[M]. Math. Oper. Statist, 1997, 11(3):403-413.

[3] Kiefer J. General Equivalence Theory for Optimum Designs (Approximate Theory)[J]. Annals of Statistics, 1974, 2(5):849-879.

[4] Dette. -H. Designing experiments with respect to “standardized” optimality criteria[J]. Journal of the Royal Statistical Society. Series B (Methodological), 1997, 59(1):97-110.

[5] 赵洪雅,关颖男,韩大伟. 二阶可加混料模型的 R-最优设计[J]. 东北大学学报,2001,(02):222-225.

[6] 孙超. 随机系数回归模型的 I_L-最优和 R-最优设计[D].上海师范大学,2012.

[7] Liu,X.,Yue,R.-X. and Kashinath Chatterjee. A note on R-optimal designs formulti-factor models[J]. J.Statist.Plann.Inference, 2014, 146(1):139-144.

[8] Liu,X. ,Yue,R.-X. and Kashinath Chatterjee. R-optimal designs in random coefficient regression models[J]. Statist.Probab.Letters.2014, 88:127-132.

[9] Liu,X. ,Yue,R.-X. and Kashinath Chatterjee. Model-robust R-optimal designs in linear regression models[J]. J.Statist.Plann.Inference, 2015, 167:135-143.

[10] Liu,X. ,Yue,R.-X. ,Xu,J. and Kashinath Chatterjee. Algorithmic construction of R-optimal designs for second-order response surface models[J]. J.Statist.Plann.Inference, 2016, 178,61-69.

[11] 徐靖. 二次响应曲面模型的 R-最优设计算法[D].上海师范大学,2016.

[12] King J, Wong W K. Minimax D-Optimal Designs for the Logistic Model[J]. Biometrics, 2000, 56(4):1263.

[13] Torsney B, Gunduz N. On optimal designs for high dimensional binary regression models[M].Optimum Design 2000, 2000, 51:275-285.

[14] Linda M. H., Ga¨ etan Kabera, Principal Ndlovu and Timothy E.O’Brien. D-optimal Designs for Logistic Regression in Two Variables[M]. mODa 8 - Advances in Model-Oriented Design and Analysis. Physica-Verlag HD, 2007, 91-98.

[15] Habib Jafari, Soliman Khazai, Yazdan Khaki and Tohid Jafari. D-optimal Designs for Logistic Regression Model with Three Independent Variables[J].Journal of Asian Scientific Research,2014, 4(3):120-124.

[16] H. T. Abebe ,F. E. S. Tan ,G. J. P. Van Breukelen, et al. Robustness of Bayesian D-optimal design for the logistic mixed model against misspecification of autocorrelation[J]. Computational Statistics, 2014, 29(6):1667-1690.

[17] 李长平, 职心乐, 刘晓红, 崔壮, 魏风江, 柯慧, 李妍, 马骏. AIC 结合最优子集法构建 logistic 回归模型在预测 2 型糖尿病并发末梢神经病变中的应用[J]. 中国卫生统计, 2010,(06):594-597+599.

[18] 陈晓兰, 任萍. 基于 Logistic 混合模型的企业信用风险评价研究[J]. 山东财政学院学报,2011,(02):90-93.

[19] 张乐勤, 陈发奎. 基于 Logistic 模型的中国城镇化演进对耕地影响前景预测及分析[J]. 农业工程学报, 2014,(04):1-11.

[20] Cheng C S. Optimal regression designs under random block-effects models[J]. Statistica Sinica, 1995, 5(2):485-497.

[21] Silvio S. Zocchi and Anthony C. Atkinson. Optimum Experimental Designs for Multinomial Logistic Models[J]. Biometrics, 1995, 55:437-444.

[22] Tan FES, Berger MPF. Optimal allocation of time points for random effects models[J]. Commun Stat Simul Comput, 1999, 28(2):517–540.

[23] Mario J.N.M. Ouwens, Frans E.S. Tan and Martijn P.F. Berger(2006). A maximin criterion for the logistic random intercept model with covariates[J]. Journal of Statistical Planning and Inference , 2006, 136(3):962-981.

[24] Tekle F B, Tan F E S, Berger M P F. Maximin D-optimal designs for binary longitudinal responses[J]. Computational Statistics & Data Analysis, 2008, 52(12):5253-5262.

[25] Debusho L K, Haines L M. V - and D -optimal population designs for the simple linear regression model with a random intercept term[J]. Journal of Statistical Planning & Inference, 138(4):1116-1130.

[26] C. Tommasi, J.M. Rodríguez-Díaz and M.T. Santos-Martín. Integral approximations for computing optimum designs in random effects logistic regression models[J]. Elsevier Science Publishers B. V., 2014, 71:1208-1220.

[27] Maram P P, Jafari H. Bayesian D-optimal design for logistic regression model with exponential distribution for random intercept[J]. Journal of Statistical Computation & Simulation, 2016, 86(10):1856-1868.

[28] 程靖，岳荣先，刘欣. 两变量随机截距模型的最优设计[J]. 数理统计与管理，2011,(03):504-511.

[29] 程靖，岳荣先. 两变量随机系数回归模型的最优设计[J]. 应用概率统计，2012，(03)：225-234.

[30] 周晓东. 线性混合效应模型的最优设计与稳健设计[D]. 上海师范大学, 2012.

[31] Jing Cheng,Rong-Xian Yue,Xin Liu. Optimal Designs for Random Coefficient Regression Models with Heteroscedastic Errors[J]. Communications in Statistics - Theory and Methods, 2013, 42:2798-2809.

[32] Li G, Majumdar D. D-optimal designs for logistic models with three and four parameters[J]. Journal of Statistical Planning & Inference, 2008, 138(7):1950-1959.

[33] Adewale A J, Wiens D P. Robust designs for misspecified logistic models[J]. Journal of Statistical Planning & Inference, 2009, 139(1):3-15.

[34] Mok M S, Sohn S Y, Ju Y H. Random effects logistic regression model for anomaly detection[J]. Expert Systems with Applications, 2010, 37(10):7162-7166.

[35] Rao C R. Linear Statistical Inference and Its Applications. Wiley, New York[J]. Mathematical Gazette, 1970, 54(388).

[36] Garcia T P, Ma Y. Optimal Estimator for Logistic Model with Distribution‐free Random Intercept[J]. Scandinavian Journal of Statistics, 2016, 43(1):156-171.

[37] A. Atkinson and A. Donove. Optimum experimental designs[M], with SAS. Oxford University Press, 2007.

[38] 方开泰，刘民千，周永道. 试验设计与建模[M]，高等教育出版社，2011.

[39] Min Yang. A-optimal designs for generalized linear model with two parameters[J]. Journal of Statistical Planning and Inference. 2008, 138(3), :624-641.

[40] Lesperance M, Saab R, Neuhaus J. Nonparametric estimation of the mixing distribution in logistic regression mixed models with random intercepts and slopes[J]. Computational Statistics & Data Analysis, 2014, 71(3):211-219.

[41] 茆诗松，丁元，周纪芗，吕乃刚，回归分析及其试验设计[M]，华东师范大学出版社，1986.

致谢

一转眼我在上海师范大学两年的研究生学习即将结束，回首两年的求学历程，对那些引导我、帮助我、激励我的人，我心中充满了感激。

首先特别感谢我的导师岳荣先教授在这两年给予我的关怀和教诲，尽管岳老师工作繁忙，但是在论文题目的选定，参考文献的检索以及论文写作的修改上，岳老师都给予了我很多关键性的指导和建议，对我的论文完成提供了很大的帮助。岳老师为人亲和，治学严谨，学识渊博，短短两年的相处令我受益终身，能师从岳老师，我的内心充满自豪。

其次还要感谢刘吉彩老师、房云老师、崔百胜老师、王周伟老师和数理学院以及商学院其他各位老师对我的教导和关心，老师们为我的学业倾注了大量心血，让我学习到很多统计知识，在此谨向各位老师表示我最诚挚的敬意和感谢！

同时我也要感谢师兄贺磊在论文写作过程中给予我的帮助，感谢我的舍友刘烨，彭钰，我们朝夕相处，互相帮助，感谢我的同学王佳燕，夏小健等，同窗之谊，我将终生难忘！

感谢父亲和母亲对我的辛苦养育，你们是我十多年的求学路上的坚强后盾，希望在不远未来我可以凭借自身的努力报答你们。

最后，衷心感谢在百忙之中抽出时间评审本文的各位学者专家！

附录

附录 1 固定截距 R-最优设计迭代程序

```
###########单因素#################################
rm(list=ls(all=TRUE))
library("nloptr")
n<-2   ##清空缓存
a<-0; b<-1;
l<-1##  系数初值
f<-function(x) c(1,x) ##  几个变量
v<-function(x) exp(1+l*x)/(1+exp(1+l*x))^2
xi<-c(0,1)##  初值
xp<-rep(1/length(xi),length(xi)) ##  重复多少次即权重 W
funM<-function(x,p){    ##  信息矩阵
   valM<-matrix(0,nrow = n,ncol = n)
   for (i in 1:length(xi)) {
      valM<-p[i]*v(x[i])*f(x[i])%*%t(f(x[i]))+valM
   }
   return(valM)
}

phix<-function(x) {      ##  敏感性函数
   val<-0
   MM<-sM%*%f(x)%*%t(f(x))%*%sM
   for(i in 1:n){
      val<-(v(x)*MM[i,i])/sM[i,i]+val
   }
   return(n-val)
}

k<-2   ##  下面寻找准则函数的全局最优的最小值
repeat{
   sM<-solve(funM(xi,xp),tol=1e-40)
   xmin <- mlsl(x0 = c(0.1), phix, lower = rep(a,1), upper = rep(b,1),
                        nl.info = F,control=list(xtol_rel=1e-8,maxeval=1000))
   valm<-signif(xmin$par,4)   ##限制小数点后四位
   if(phix(b)<phix(valm)) valm<-b
   if(phix(a)<phix(valm)) valm<-a
   dimv<-max(which(xi<=valm))
   if(valm==b) dimv<-dimv-1
   #dimv<-max(which(xi<=valm))
```

```
   if(abs(xi[dimv]-valm)<0.005){
      xp<-xp*(1-1/k)
      xp[dimv]<-1/k+xp[dimv]
   } else if(abs(xi[dimv+1]-valm)<0.005){
      xp<-xp*(1-1/k)
      xp[dimv+1]<-1/k+xp[dimv+1]
   } else{
      xi<-c(xi[1:dimv],valm,xi[(dimv+1):length(xi)])
      xp<-c(xp[1:dimv]*(1-1/k),1/k,xp[(dimv+1):length(xp)]*(1-1/k))
   }
   k<-k+1
   if(k>5000) break     ##循环 5000 次
}
k

##result   作图 x 与准则函数 y
#xi<-c(0,1)
#xp<-c(0.5731,0.4269)
sM<-solve(funM(xi,xp),tol=1e-40)
x<-seq(a,b,by=0.01)
y<-numeric(length(x))
for(i in 1:length(x)) y[i]<-n-phix(x[i])
max(y[i])
plot(x,y,type = "l")
lines(x,rep(0,length(x)))
abline(h=2,lwd=1)
N<-which(xp>=0.001)
title("参数为 1")
xi[N]
xp[N]

##效率计算
rm(list=ls(all=TRUE))
l<-2.5
f<-function(x) c(1,x) ##  几个变量
v<-function(x) exp(1+l*x)/(1+exp(1+l*x))^2
n<-2
xi<-c(0,1)##  初值
xp<-rep(1/length(xi),length(xi))

xi<-c(0,0.9576,0.9642)
xp<-c(0.5559,0.0578,0.3858)
```

```
funM<-function(x,p){   ## 信息矩阵
   valM<-matrix(0,nrow = n,ncol = n)
   for (i in 1:length(xi)) {
      valM<-p[i]*v(x[i])*f(x[i])%*%t(f(x[i]))+valM
   }
   return(valM)
}
sM1<-solve(funM(xi,xp),tol=1e-40)
sM1[1,1]*sM1[2,2]
v(xi)
#########双因素##############################################
rm(list=ls(all=TRUE))
library("nloptr")
n<-3
a<-0; b<-2
l1<-2;l2<-2
f<-function(x1,x2) c(1,x1,x2)
v<-function(x1,x2) exp(1+l1*x1+l2*x2)/(1+exp(1+l1*x1+l2*x2))^2
x1.i<-seq(a,b,by=1);   x2.i<-seq(a,b,by=1)
xi<-merge(x1.i,x2.i)
xp<-rep(1/nrow(xi),nrow(xi))
funM<-function(x,p){
   valM<-matrix(0,nrow = n,ncol = n)
   for (i in 1:nrow(xi)) {
         valM<-p[i]*v(xi[i,1],xi[i,2])*f(xi[i,1],xi[i,2])%*%
               t(f(xi[i,1],xi[i,2]))+valM
   }
return(valM)
}
phix<-function(x) { # x is a vector of 2 degree
   val<-0
   MM<-sM%*%f(x[1],x[2])%*%t(f(x[1],x[2]))%*%sM
   for(i in 1:n){
      val<-MM[i,i]/sM[i,i]+val
   }
   return(n-v(x[1],x[2])*val)
}

k<-2
repeat{
   sM<-solve(funM(xi,xp),tol=1e-40)
   xmin <- mlsl(x0 = c(0.3,0.3), phix, lower = rep(a,2), upper = rep(b,2),
               nl.info = F,control=list(xtol_rel=1e-8,maxeval=1000))
```

```
   valm<-signif(xmin$par,4)
   if(phix(c(b,b))<phix(valm)) valm<-c(b,b)
   if(phix(c(a,b))<phix(valm)) valm<-c(a,b)
   if(phix(c(b,a))<phix(valm)) valm<-c(b,a)
   if(phix(c(a,a))<phix(valm)) valm<-c(a,a)
   #dimv<-max(which(xi[,1]<=valm[1] & xi[,2]<=valm[2]))
   dimv1<-max(which(x1.i<=valm[1]))
      if(valm[1]==b) dimv1<-dimv1-1
   dimv2<-max(which(x2.i<=valm[2]))
      if(valm[2]==b) dimv2<-dimv2-1
#####################################
   if(abs(x1.i[dimv1]-valm[1])<0.005){
      valm[1]<-x1.i[dimv1]
      } else if(abs(x1.i[dimv1+1]-valm[1])<0.005){
      valm[1]<-x1.i[dimv1+1]
      } else{
      x1.i<-c(x1.i[1:dimv1],valm[1],x1.i[(dimv1+1):length(x1.i)])
      }
#####################################
   if(abs(x2.i[dimv2]-valm[2])<0.005){
      valm[2]<-x2.i[dimv2]
   } else if(abs(x2.i[dimv2+1]-valm[2])<0.005){
      valm[2]<-x2.i[dimv2+1]
   } else{
      x2.i<-c(x2.i[1:dimv2],valm[2],x2.i[(dimv2+1):length(x2.i)])
   }
#####################################
   dimv<-which(xi[,1]==valm[1] & xi[,2]==valm[2])
   if(length(dimv)==1){
      xp<-xp*(1-1/k)
      xp[dimv]<-1/k+xp[dimv]
   } else if(length(dimv)==0){
      xi<-rbind(xi,valm)
      xp<-c(xp*(1-1/k),1/k)
   }
   k<-k+1
   if(k>5000) break
}
N<-which(xp>=0.001)
xi[N,]
xp[N]

library(rgl)
```

```
sM<-solve(funM(xi,xp),tol=1e-40)
x1<-seq(0,2,by=0.05)
x2<-seq(0,2,by=0.05)
x<-merge(x1,x2)

y<-numeric(length(x1))
for(i in 1:length(x1)) {
      q<-c(x[i,1],x[i,2])
      y[i]<-n-phix(q)
}
plot3d(x1,x2,y,type ="l",size=3)
max(y[i])

##效率计算
rm(list=ls(all=TRUE))
l1<-2;l2<-2
a<-0; b<-2
f<-function(x1,x2) c(1,x1,x2)
v<-function(x1,x2) exp(1+l1*x1+l2*x2)/(1+exp(1+l1*x1+l2*x2))^2
n<-3
#x1.i<-seq(a,b,by=1);    x2.i<-seq(a,b,by=1)
#xi<-merge(x1.i,x2.i)
#xp<-rep(1/nrow(xi),nrow(xi))
x1<-c(0,1.198,0,1.192,0);x2<-c(0,0,1.198,0,1.192)
xi<-cbind(x1,x2)
xp<-c(0.404,0.0738,0.073,0.2236,0.224)

funM<-function(x,p){
   valM<-matrix(0,nrow = n,ncol = n)
   for (i in 1:nrow(xi)) {
         valM<-p[i]*v(xi[i,1],xi[i,2])*f(xi[i,1],xi[i,2])%*%
               t(f(xi[i,1],xi[i,2]))+valM
   }
return(valM)
}
sM1<-solve(funM(xi,xp),tol=1e-40)
sM1[1,1]*sM1[2,2]*sM1[3,3]
```

附录 2 随机截距 R-最优设计迭代程序

```
############随机截距单因素#################################
rm(list=ls(all=TRUE))
library("nloptr")
n<-2   ##清空缓存
a<-0; b<-1;
l<-2;sd<-1##  系数初值
xi<-c(0,1)
ei<-rnorm(2,mean=0,sd=1)
f<-function(x) c(1,x)
v<-function(x) {
     w<-c()
     for (i in 1:length(x)){
     we<-(exp(1+l*x[i]+ei[i])/(1+exp(1+l*x[i]+ei[i]))^2)
     w<-c(w,we)
}
return(w)
}

#xi<-cbind(x1.i,x2.i)
xp<-rep(1/length(xi),length(xi))
funM<-function(x,p){   ##  信息矩阵
   valM<-matrix(0,nrow = n,ncol = n)
   for (i in 1:length(xi)) {
      valM<-p[i]*v(x[i])*f(x[i])%*%t(f(x[i]))+valM
   }
   return(valM)
}
phix<-function(x) {     ##  敏感性函数
   val<-0
   MM<-sM%*%f(x)%*%t(f(x))%*%sM
   for(i in 1:n){
      val<-(v(x)*MM[i,i])/sM[i,i]+val
   }
   return(n-val)
}
k<-2   ##  下面寻找准则函数的全局最优的最小值
repeat{
   sM<-solve(funM(xi,xp),tol=1e-40)
   xmin <- mlsl(x0 = c(0.1), phix, lower = rep(a,1), upper = rep(b,1),
                        nl.info = F,control=list(xtol_rel=1e-8,maxeval=1000))
   valm<-signif(xmin$par,4)   ##限制小数点后四位
   if(phix(b)<phix(valm)) valm<-b
```

```
   if(phix(a)<phix(valm)) valm<-a
   dimv<-max(which(xi<=valm))
   if(valm==b) dimv<-dimv-1
   #dimv<-max(which(xi<=valm))
   if(abs(xi[dimv]-valm)<0.005){
      xp<-xp*(1-1/k)
      xp[dimv]<-1/k+xp[dimv]
   } else if(abs(xi[dimv+1]-valm)<0.005){
      xp<-xp*(1-1/k)
      xp[dimv+1]<-1/k+xp[dimv+1]
   } else{
      xi<-c(xi[1:dimv],valm,xi[(dimv+1):length(xi)])
      xp<-c(xp[1:dimv]*(1-1/k),1/k,xp[(dimv+1):length(xp)]*(1-1/k))
   }
   k<-k+1
   if(k>5000) break     ##循环 5000 次
}
k

##result   作图 x 与准则函数 y
#xi<-c(0,1)
#xp<-c(0.5731,0.4269)
sM<-solve(funM(xi,xp),tol=1e-40)
x<-seq(a,b,by=0.01)
y<-numeric(length(x))
for(i in 1:length(x)) y[i]<-n-phix(x[i])
max(y[i])
plot(x,y,type = "l")
lines(x,rep(0,length(x)))
abline(h=2,lwd=1)
N<-which(xp>=0.001)
title("参数为 2")
xi[N]
xp[N]

##效率计算
rm(list=ls(all=TRUE))
library("nloptr")
n<-2   ##清空缓存
a<-0; b<-1;
l<-1;##  系数初值
xi<-c(0,1)
#ei<-rnorm(2,mean=0,sd=1) ei<-c(-0.3943645,0.0239947)
```

```
#xp<-c(0.6431,0.3569)
xp<-rep(1/length(xi),length(xi))

f<-function(x) c(1,x)
v<-function(x) {
    w<-c()
    for (i in 1:length(x)){
    we<-(exp(1+l*x[i]+ei[i])/(1+exp(1+l*x[i]+ei[i]))^2)
    w<-c(w,we)
}
return(w)
}

funM<-function(x,p){    ## 信息矩阵
  valM<-matrix(0,nrow = n,ncol = n)
  for (i in 1:length(xi)) {
    valM<-p[i]*v(x[i])*f(x[i])%*%t(f(x[i]))+valM
  }
  return(valM)
}
sM1<-solve(funM(xi,xp),tol=1e-40)
sM1[1,1]*sM1[2,2]
```