C – Optimal Sequential Design on Compound Poisson Regression Model

S.Joshua David * and C. Santharam**

[1]*Department of Statistics, Loyola College, Chennai – 600034*
Tamil Nadu, India

Abstract: *C optimal design is use to analysing the geometry of the canonical design space a design point is required when estimating point of the height and response curve at the maximum or minimum to estimating β, which is related to width of the response curve. An analogous of C optimality design criterion was developed using variance covariance matrix, for GLM, We consider the problem of finding an optimal design model under a compound Poisson distribution with any number of independent variables and a reciprocal link additive linear predictor model. One of the main interests to determine the optimum operating conditions is to locate the point of maximum or minimum response on response variables to be optimized are different quality aspects of a product, such as yield or strength. The fact that the optimal design depends on the unknown true model parameters for the compound Poisson regression model with gamma link relative is very use full. One approach to solve this problem is to use sequential designs for binary data with the purpose to estimate a percentile of the response curve are treated. The theory of stochastic approximation provides an alternative method for estimation of the optimum point of a response function. The design points are determined successively will converge to the point of optimum response.*
Key words: *Compound Poisson, Locally C optimal design, sequential design, stochstic link model*

I.   Introduction

The optimum operating conditions to locate the point of maximum or minimum response on response variables to be optimized are different quality aspects of a product, such as yield or strength. The fact that the optimal design depends on the unknown true model parameters for the compound Poisson model, there would be no point in performing experiments if the true parameters were known. One approach to solve this problem is to use sequential designs. Starting with a preliminary estimate or guess of the parameters a locally optimal design can be constructed. This design is then used to update the parameter estimates, which in turn leads to another locally optimal design, new parameter estimates and so on. Sequential designs for sequential data with the purpose to estimate a percentile of the response curve are treated in   Wu (1985). The theory of stochastic approximation provides an alternative method for estimation of the optimum point of a response function. Stochastic approximation started with the work of Robbins and Monro (1951) and Kiefer and Wolfowitz (1952) and is a nonparametric sequential approach. The design points are determined successively according to a recursive scheme such that the resulting sequence will converge to the point of optimum response.   A stochastic approximation method is evaluated for binary data in the case of estimating a percentile of the response curve in Wu (1985). Wu (1986) goes through the connections and differences between the stochastic approximation method and a parametric approach based on maximum likelihood estimation. The parametric approaches of optimal designs are derived sequentially based on the assumption that the most recent parameter estimates are true.

Sequential Designs

Locally C-optimal designs consisting of two equally weighted designs points are computed at each stage, using two versions of the standardized information matrix of the weighted version that takes into account the information of the observed points. In accordance with a recursive scheme successive observations are made on the response variable in a way that the sequence of design points $x_1, x_2 \ldots$ converges to μ. This approach is adaptive in the sense that it makes use of the information gained so far by adjusting the step from x_r to $x_r + 1$. The second approach is parametric; optimal designs are derived sequentially based on the assumption that the most recent parameter estimates are true. Locally c-optimal designs consisting of two equally weighted designs points are computed at each signs consisting of two equally weighted designs points are computed at each stage, using two versions of the standardized information matrix. The first is the regular version that reflects the information matrix. The first is the regular version that reflects the information in the candidate design and the second one is a weighted version that also takes into account the information in the already observed points.

Let f (x) denote a response function that is unknown. Robbins and Monro (1951) give a stochastic approximation method for finding the solution x=θ to the equation f (x)=c, where c is a constant. It is assumed that f(x)=c has a unique solution θand that for every observations con be made on a random variable Y(x) such that E(Y(x))=f(x).. Starting at an arbitrary x_1 consecutive observations are made on Y(x) at x_2, x_3…. in such a way that x_r converges to as θ as $r \rightarrow \infty$. At x_r the next design point, x_r+1, is chosen according to the following scheme.

$$x_r+1 = x_r + \beta_r(y_r - d),$$

where $\{\beta_r\}$ is a fixed infinite sequence of decreasing positive constants satisfying $\sum \beta_r^2 < \infty$. The choice of the constants $\{\beta_r\}$ is important for the performance of the sequence $\{x_r\}$. In the sense of achieving minimal asymptotic variance it is optimal to set β_r equal to $[rf'(\theta)]'$, see e.g. Chung (1954) and Sacks (1958). However, $f'(\theta)$ f is generally unknown and needs to be estimated. One possible estimator is the least squares estimator of the slope in the regression Y on x. An adaptive version of the Robbins-Monro procedure, where the estimate of $f'(\theta)$ is updated at each step by using the least squares estimator, is given by

$$x_r+1 = x_r - (\eta\tilde{\lambda}_r)^{-1}(y_r - d),$$
$$\text{where } \tilde{\lambda}_r = \frac{\sum(y_i x_i - \bar{x}_r)}{\sum(x_i - \bar{x}_r)^2} , \bar{x}_r = \frac{\sum x_r}{n}$$

This procedure is described by Anbar (1978) and was reviewed and evaluated for binary data in Wu (1985). The adaptive procedure was proven to be asymptotically equivalent to the non adaptive Robbins-Monro procedure with $\beta_r = [rf'(\theta)]^{-1}$ by Anbar (1978) and Lai and Robbins (1979).

The Robbins-Monro procedure was further developed by Kiefer and Wolfowitz (1952) and extended to the case of estimation the maximum at the point θ and f(x) is assumed to be strictly increasing (decreasing) for x < θ (x > θ). Furthermore it is assumed that observations can be taken on the random variable Y(x) at any level x and that E [Y (x)]= f (x). The principle is the same as for the Robbins-Montro procedure, that is successive observations are made on Y(z) according to a specified scheme. The principle is the same as for the Robbins-Monro procedure, that is successive observations are made on y(x) according to a specified scheme.

The difference is that observations are made in pairs at each step. Starting at an arbitrary x_1, the following $x_2, x_3,\ldots$ are obtained by making observations at $x_r \pm D_r$. i.e. both $Y(x_r-D_r)$ are observed at each step. The sequence $\{x_r\}$ is defined as

$$x_r+1 - x_r + \beta_r \frac{y(x_r + D_r) - y(x_r - D)}{2D_r} = x_+ \beta_r z_r$$

and converges to θ as $\rightarrow \infty$ $\{a_r\}$ and $\{D_r\}$ are pre assigned infinite sequences of positive numbers such that

$$D_r \longrightarrow 0,$$
$$\sum \beta_r = \infty$$
$$\sum \beta_r D_r < \infty$$
$$\sum \beta_r^2 D_r^2 < \infty$$

For example $\beta_r = r^{-1}$ and $D_r = r^{-1/3}$ satisfy these conditions. The random variable $Z_r = (Y(x_r + D_r) - [Y(x_r - D_r)]/(2D_r)$ can be viewed as an approximation to the derivative of the response function at x_r. This reduces the problem to find the maximum of f(x) to that of finding the solution to the equation $f'(x) = 0$. The recursion can then be thought of as a special case of the Robbins-Monro method for finding the solution to $f'(z) = 0$. by making successive observations on Z.

In this thesis the procedure of Kiefer and Wolfowitz is combined with the adaptive Robbins-Monro procedure (5.2) described above for estimating the maximum of the response function π(x) , abbreviated as KW hereafter. At each step observations are taken at $x_r \pm C_r$ so that the numerical derivative becomes $Z_r = [Y(x_r + D_r) - Y(x_r + D_r)]/(2mD_r)$ A graphical illustration is given in Figure Starting at an arbitrary x_1, the value on x_r, is then updated via the adaptive Robbins-Monro procedure

$$x_r+1 = x_r - \eta\tilde{\lambda}_r^{-1}z_r$$

where $\tilde{\lambda}$ is the least squares estimator of the slope in the regression Z on x. The estimate of the optimum point at the r:th step will then be given by

$$= \tilde{\theta}_r xe.$$

Kiefer and Wolfowitz (1952) put some conditions on the response function that prevents it from being too steep or to flatten out towards zero. If the curve is too steep it may cause unduly large changes in x and any observations taken where the response curve is zero will be uninformative, making it impossible to know in which direction to take the next step. Kiefer and Wolfowitz comment however, that it will be sufficient if the conditions are fulfilled in an interval (D_1, D_2). No observations will be taken outside the limits of this interval. The flatness of the logistic response curve may pose some problems unless there is knowledge about an appropriate interval before the experiment is started. In some situations it may be the case that such information is no available.

The success of this procedure is also dependent upon good choices of the starting values r_1, λ_1 and D_1. If the starting point is too far off, there will be essentially zero probability of obtaining a response so that no

information can be gained. The starting value for D_r should be small enough to avoid unduly large changes in x and large enough to enable the sequence to move away from a bad starting point. The relation between the number of observations taken at each step and the number of steps is another aspect that probably will be important,
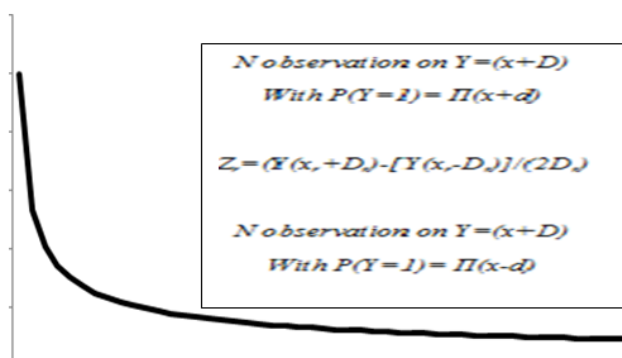
C-optimal Sequential Designs

The point of maximum response of the quadratic response curve is equal to the parameter . Estimating with minimum variance would be desirable justifying the use of a c-optimal design with $c = (-1,0,1)'$. As an attempt to handle the problem of parameter dependence, the c-optimal design is updated sequential. It was established in Chapter 3 that a c-optimal design for estimation of the point $x= 1/\mu$ consists of two points with equal weight. In essence, two points are taken at a certain distance from each other, a distance that does not necessarily decrease, as opposed to the KW approach. This parametric sequential approach can be described by the following steps:

1. Choose an initial design ξ (initial)
2. Estimate the parameters $\tilde{\theta}$ (initial)
3. Minimize $\tilde{\theta} M^{-1}$ to find the locally c-optimal design given
4. Take m observations at the design points ξ^{*1} in an estimate the parameters: $\tilde{\theta}^{(1)}$
5. Maximize $\tilde{C}^T M^{-1}\left(\xi, \tilde{\theta}^{(1)}\right) \tilde{c}$ to find the locally c-optimal design given $\tilde{\theta}^{(1)} \longrightarrow \xi_{*2 =}\begin{Bmatrix} x_1 & x_2 \\ 0.5 & 0.05 \end{Bmatrix}$ etc.

The parameter estimates $\tilde{\theta}$ are the usual maximum likelihood estimates. the ewstimate of the optimum point afater r steps is then given by $\tilde{\mu}$.

If the initial design contains N* observations and 2m observations are taken at each step there are N*+2mr=N_{tat} observations in total after r steps. The observed standardized information matrix can then be expressed as

$$\text{Mobs}\left(\xi\tilde{\theta}\right) = \sum_{i = 1}^{N\,lmt} \frac{1}{N_{tat}}\widetilde{\pi}_i \begin{pmatrix} -1 \\ -(x_{i-}\widetilde{\mu})^2 \\ 2\tilde{\beta}\,(x_{i-}\tilde{\mu}) \end{pmatrix}' \begin{pmatrix} -1 \\ -(x_{i-}\widetilde{\mu})^2 \\ \widetilde{2\beta}\,(x_{i-}\tilde{\mu}) \end{pmatrix}$$

where $\widetilde{\pi}_{i=}\left(e^{x_i\,\tilde{\theta}}\right)/\left(1 + e^{x_i\,\tilde{\theta}}\right)$ Another version of the previous sequential design, called model 2, makes use of the observed information matrix in the following way. $M\left(\xi\tilde{\theta}\right)$ is replaced by a weighted information matrix computed as

$$M_w\left(\xi\tilde{\theta}\right) = \frac{N_{tat}}{N_{tat}+2m} M_{obs}\left(\xi_{obs}\,\tilde{\theta}\right)+ \frac{2m}{N_{tat}+2m} M\left(\xi\tilde{\theta}\right).$$

and $\tilde{c}^T M^{-1}{}_W\left(\xi\tilde{\theta}\right)\tilde{c}$ is minimized instead. M_{obs} is the information matrix based on all observations up to this point and ξ_{obs} deign consisting of these observations. This is potentially an improvement because it also takes into account the information from the proceeding steps. Besides the changed information matrix everything is the same as for the parametric sequential procedure.

Replication Study of Sequential

The results of the simulations are also reported in Fackle Fornius (2008). There are many questions to bring clarity about including: it's an approaches superior? Is there an optimal choice of c for the approach? Is the weighted version of the information matrix preferred over the unweighted? What is best, taking few steps with many observations at each step, or taking many steps with few observations? Are the performances robust to misspecifications of the model?

Two parameter sets are considered here: θ_P= (-1, 0.1, 1)'. The sequential design approaches are analyzed in the simulations. Different sample sizes ranging from N=50 to N=100 as well as different batches

sizes are explored. To study the effects of specifications in the model an alternative to the linear predictor is also tested.

All the three approaches are started from the same initial design before they take separate paths. Two different initial designs are used. Design 1 is constructed to be better than Design 2 for estimation of the point of maximum response μ. It is better in the sense that it is symmetric around μ = 0. C-optimal designs are similarly characterized by design points placed down words sloping but limited to two points. However, it is not possible to obtain maximum likelihood estimates of the model parameters with a C-optimal design because two points are not sufficient to estimate three parameters. Design 2 is constructed based on the same principle but pushed aside, intended to represent a beforehand assumption that is close to 1. The initial designs consist of a fairly large amount of observations (N=50) to avoid the problem of non-existing maximum likelihood estimates, which is also a motive for choosing many design points.

The C-optimal procedures are made on the response variable Y at the design points specified by the initial design. Initial maximum likelihood estimates $\tilde{\theta}^{(initial)}$ are then calculated. Both approaches are now ready to start from Step 3 with the minimization of $\tilde{c}^T M^{-1} \tilde{\theta}^{(initial)}$. m response values are now generated at each of the two c-optimal design points, resulting in a new C-optimal design and so on.

The Kiefer and Wolfwitz approach needs at least two observations on

$$Z = [Y(x+c) + Y(z-c)] / (2mc)]$$

before a least squares estimate of the slope parameter in the Z on x regression can be obtained. As discussed before the Kiefer and Wolfwitz approach is so sensitive to the choices of x_1 and β_1 Some effort is therefore put in selecting good starting values. A preliminary estimate of the point of maximum response is obtained from the initial design. The first design point in Kiefer and Wolfwitz is set equal to this estimate:

$$x_1 = \tilde{\mu}_{initial}$$

Observations are made at these three design points that will result in three observation on Z which then can be used to obtain the starting value β_1 For different design points are obtained from Several values on D_r are tested.

Table: Support points of the locally C -optimal design for the Compound Poisson model for various values of μ_i and various values of N and also varying β

Mean Value	Constant Value		Initial Design point	Model 1	Model 2	Model 3
0.990099	β=.1	N = 50	0.953463			
	$\beta_0=1$			0.426401		
	$\beta_0=0$				0.596159	
	$\beta_0=-1$					1.025978
0.952381	β=.3	N = 50	0.632456			
	$\beta_0=1$			0.254		
	$\beta_0=0$				0.282843	0.302199
	$\beta_0=-1$					
0.900901	β=.7	N=100	0.339032			
	$\beta_0=1$			0.181071		
	$\beta_0=0$				0.190693	
	$\beta_0=-1$					0.196305
0.854701	β=.9	N=100	0.247689			
	$\beta_0=1$			0.14825		
	$\beta_0=0$				0.153393	
	$\beta_0=-1$					0.156269

Sequential Design Points
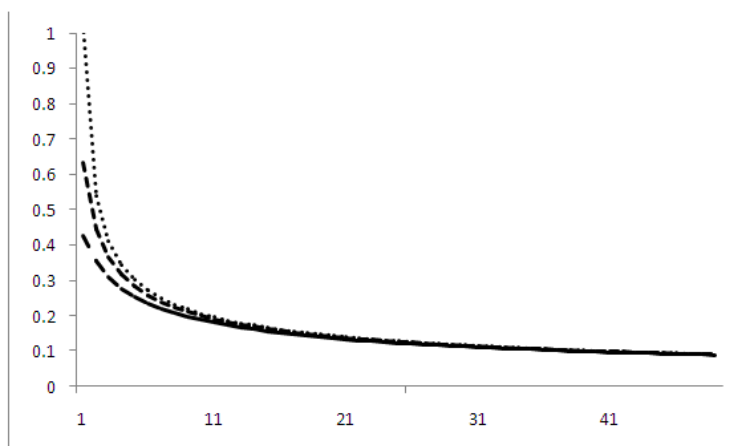
The above graph is to identify the location sequential optimal points using inverse link relative model. The points are potentially critical points of the model parameters especially estimation of the points of maximum or minimum probability D(x). The graph shows the points of intersection between Z values the observation attaints optimum point to study the height of the response curve in inverse link GLM model optimal design is decreasing when the mean values are decreasing. The optimal design has equal allocation between [-1, 1].

II. Conclusion

The number of design points in a C optimal sequential design points for estimating the model parameter varies depending on which parameter to be estimated. The analytic contraction of C -optimal sequential Compound Poisson model with inverse link function were derived to obtain locally C -optimal design, the solution to these formulae can be obtain numerically using R- software in contrast to using for examples algorithms which involve issues like finding a good initial design. Thus the C-optimal sequential design depends on the various parameters of β with varying $\Box(x)$ values it is the function of mean values. We conclude that when the optimal designs with inverse link function have a linear functional relationship between function of mean values and the various predicted vales. When the function of mean values is decreasing the predictors are decreasing the design involves asymptotic information matrix.

The linear predictors of inverse link function attain maximal optimal values for the different β values it converges to zero for various functional mean values. When the varying values of β respected models attain exact limit point if the β_0 decreases the inverse link predictor of sequential design attain maximum value point.

References:

[1]. Silvey, D., 1980. Optimal Design, Chapman and Hall, London.
[2]. Bai, D.S., Chung, S.W.: An optimal design of accelerated life test for exponential distribution.Reliab. Eng. Syst. Saf. 31, 57–64 (1991).
[3]. Hugh A. Chipman and William J. Welch. D-optimal Design for Generalized Linear Models, Chicago (1996).
[4]. Holger Dette, Viatcheslav B. Melas and Weng Kee Wong: Locally D-Optimal Design for Exponential Regression Models. Los Angeles. Statistica Sinica 16(2006).
[5]. Santharam, C., Ponnusamy, K.N and Chandrasekar, B.,Universal optimality of nearest neighbour balanced block designs using ARMA models. Biometrical.J. 32,725 – 730, 1996.
[6]. Rubanraj. B, Santharam. C and Ramesh Kumar, (2012) MV-Optimality of Nearest Neighbour Balanced Block Designs using First order and Second order correlated models, International journal of Statistika and Mathematika, (ISSN: 2277-2790 E-ISSN: 2249-8605)
[7]. Senthil kumar.R, C. Santharam (2012), Efficiency of Nearest Neighbour Balanced Block Designs for Correlated Observations (ARMA Models), International Journal of Statistika and Mathematika, ISSN: 2277- 2790 E-ISSN: 2249-8605,Volume 4, Issue 1, 2012 pp 01-05.
[8]. Senthil kumar. R, C. Santharam (2012), Efficiency of Nearest Neighbour Balanced Block Designs for Correlated Observation, International Journal of Statistika and Mathematika, ISSN: 2277- 2790 E-ISSN: 2249-8605, Volume 3, Issue 3, 2012 pp 115-120.
Joshua David. S, C. Santharam (2012), D-Optimal Design for Compound Poisson Regression Model, International Journal of Statistika and Mathematika, ISSN: 2277- 2790 E-ISSN: 2249-8605, Volume 4, Issue 2, 2012 pp 47-49.
[9]. Joshua David. S, C. Santharam (2013), C-Optimal Design for Compound Poisson Regression Model, International Journal of Statistics and Analysis.ISSN 2248-9959 Volume 3, Number 5 (2013), pp. 457-464
[10]. Jeevitha. M, C. Santharam (2013), optimality Efficiency of Circular Neighbour Balanced Block Designs for Correlated Observation, International Journal of Statistika and Mathematika, ISSN: 2277- 2790 E-ISSN: 2249-8605, Volume 7, Issue 2, 2013 pp 19 - 23.
Ruban Raj. B, C. Santharam (2013), MV – optimality of Nearest Neighbour Balanced Block Design (NNBD)using First Order Correlated Models for Five Treatments, International Journal of Statistics and Analysis, ISSN 2248 – 9959, Volume 3, Issue 4, 2013 pp 379-384.
[11]. Senthil Kumar. R, C. Santharam. (2013), Efficiency of NNBD over NNBIBD using First Order Correlated Models, IOSR Journal of Mathematics (IOSR-JM), e-ISSN: 2278-5728, p-ISSN: 2319- 765X, Volume 7, Issue 5 (Jul. – Aug. 2013), PP 06-13.
[12]. Senthil Kumar. R, C. Santharam. (2013), Efficiency of Nearest Neighbour Balanced Block Designs using ARMA models, International Journal of Statistics and Systems, ISSN: 0973-2675 Volume 8, Number 1, pp. 59-71.
[13]. Joanna Kezia. E and JoshuaDavid. S (2013), The Effect of Organic and Inorganic Fertilizers on the Growth of Radish, IOSR Journal Of Humanities And Social Science (IOSR-JHSS)   Volume X, Issue X (Nov. - Dec. 2013), PP 01-00. e-ISSN: 2279-0837, p-ISSN: 2279-0845.