International Journal of Formal Sciences: Current and Future Research Trends (IJFSCFRT)
ISSN (Print) , ISSN (Online)

https://ijfscfrtjournal.isrra.org/index.php/Formal_Sciences_Journal/index

Bayesian Information Based Criteria for Poisson Regression Models

Hillary Ali[a*], Nwaosu S. C.[b], Lasisi K. E.[c], A. Abdulkadir[d]

[a]*Department of Mathematics, University of Jos, Nigeria*
[b]*Department of Statistics, Joseph Sawuaan Tarka University of Agriculture, Makurdi, Nigeria*
[c,d]*Department of Mathematical Sciences, Abubakar Tafawa Balewa University,P. M. B. 0248, Bauchi, Nigeria*
[a]*Email: auhills1@gmail.com*, [b]*Email: alih@unijos.edu.ng*
Phone Number: +2348036984835

Abstract

Bayesian optimal designs for count responses analyzed with Poisson regression describing a linear health effect are considered. To overcome the problem of dependence of Bayesian designs on the choice of prior distributions, Bayesian D & A-optimal designs are proposed for Poisson regression models. The results showed that the optimal number of time points depends on the subject-to-measurement. Also, Bayesian D & A-optimal designs are highly efficient and robust under models considered when implementing the efficiencies of designs with the Bayesian D- & A-optimal designs in modelling students' academic performance. The best design is found in one out of ten tries.

***Keywords*:** Bayesian Poisson Regression**;** D-Optimality; A-Optimality; Efficiency.

1. Introduction

Experimental Design is a technique for the concept of a priori, taking place before data is collected, and hence the Bayesian paradigm is a particularly appropriate approach to take. Bayesian methods allow available prior information on the model to be incorporated into both the design of the experiment and the analysis of the resulting data, and produce posterior distributions that are interpretable by scientists. They also reduce reliance on unrealistic assumptions and asymptotic results that may be inappropriate for small to medium-sized experiments. The Bayesian approach to design enables realistic and coherent accounting for the substantial model and parameter uncertainties that usually exist before an experiment is performed and it is also a natural framework for sequential inference and design.

----

* Corresponding author.

An important problem where Bayesian methods can have substantial impact is optimal design for linear modelling, which relies on some prior information being available about the unknown values of the model parameters. Atkinson and his colleagues [2].

A Bayesian approach relaxes the requirement of locally optimal design criteria to specify particular values of the parameters. Fully Bayesian design, predicated on using the posterior distributions for inference, is also less reliant on the asymptotic assumptions that underpin most classical design for generalized linear models. Until very recently, optimal Bayesian design has not evolved far from the methods reviewed by Chaloner and Verdinelli [5]. Development and application of methods for Bayesian design have lagged behind the progress made in inference and modelling due to the additional complexity introduced by the need to integrate over the (as yet) unobserved responses, in addition to unknown model parameters. Hence, methodology has been restricted to simple models and fully sequential, one-point-at-a-time, procedures. Ryan and his colleagues [4].

Design of Experiments for count responses are very important in biological and clinical trials. Discussion of the non-Bayesian design for Poisson regression models can be found in Finney [2]. However, a design optimal to a best guess may not be efficient for parameter values close to the best guess so that the design is not quite robust to the parameter misspecification. Reference [6] examined the Bayesian optimal design for the one-variable Poisson regression model using the Nelder-Mead algorithm [44]. However, since the Nelder-Mead algorithm is a local-optimization method, the selection of starting design points has great influence on the performance of the procedure in getting to the global optimum. Furthermore, it would be much less efficient to use this algorithm for multi-variable nonlinear regression models. Here, the Bayesian optimal design approach is proposed for multiple Poisson regression models.

2. Materials and Methods

Bayesian D-Optimal Design for Poisson Regression Models

Poisson regression model is similar to regular multiple regression except that the dependent (*Y*) variable is an observed count that follows the Poisson distribution. Thus, the possible values of *Y* are the nonnegative integers: 0, 1, 2, 3, and so on. It is assumed that large counts are rare. Hence, Poisson regression is similar to logistic regression, which also has a discrete response variable. However, the response is not limited to specific values as it is in logistic regression. Counted data are often modeled using a Poisson model. The Poisson generalized linear model, often called the Poisson regression model, which is very useful in modelling counts, is widely used in biological and clinical experiments, assumes that *y* is Poisson with mean μ (and therefore variance μ). The link function is typically chosen to be the logarithm, so that $\log \mu = X\beta$. the fundamental Poisson regression model for an observation *i* is written as

$$\Pr\left(Y_i = y_i / \mu_i t_i\right) = \frac{e^{-\mu_i t_i}\left(\mu_i t_i\right)^{y_i}}{y_i!} \quad \ldots (1)$$

Where
$$\mu_i = t_i \mu\left(X_1^1 \beta\right) = t_i \exp\left(\beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}\right)$$

the unknown parameters are involved in the Fisher information matrix; thus, the D-optimal design is dependent on the parameter values for the Poisson regression model. In order to find and to implement the D-optimal design, the model parameter values should be known. However, in most situations, the parameter values are unknown so that a guess to the true values is needed. When the guesses are not so close to the true parameters, the design that resulted from those guesses may not be optimal. Hence the design may not be quite robust to the parameter misspecification, Wang and his colleagues [45].

Wang and his colleagues [49] presented the D-optimal design and sequential design results for Poisson regression model using the non-Bayesian design approach. However, the robustness of the non-Bayesian optimal design is a problem when the model is misspecified. Design optimality criteria can be used to select a set of experimental conditions with optimal properties. Suppose a generalized linear model is expressed as

$$E(y) = f\left(x^T \mu\right) \qquad \ldots (2)$$

Further, suppose an experiment is to be designed by choosing n values of the design variables x from an experimental region X. At each value of x_i, for $i = 1, 2, \ldots, n$, an independent observation y_i will be observed. Denote by ϑ the design measure, $y^T = \left(y_1, \ldots, y_n\right)$ the total data, $\mu^T = \left(\mu_1, \ldots, \mu_p\right)$ the unknown parameters, and $p\left(y_i / \mu, x_i\right)$ the density function of observation y_i

Under certain model assumptions and design optimality criteria, the X that maximizes the Fisher information about the unknown μ is selected. This is usually done by choosing the X to optimize a certain function of the expected Fisher information matrix,

$$I(\mu, \vartheta) = -E_y\left\{\frac{\partial}{\partial \mu^T}\left[\frac{\partial \log p(y/\mu, x)}{\partial \mu}\right]\right\} \qquad \ldots (3)$$

When the model of interest is nonlinear or when a nonlinear function of the coefficients of a linear model is of interest, the experimental design is usually more difficult to deal with, since in such cases, the information in (2) usually depends on the unknown parameters which cannot be separated as a simple multiplier. In non-Bayesian designs, it is common to replace the parameters in (2) that cannot be separated as a data-independent multiplier replaced by initial guesses. Hence when the guesses are not so close to the true parameters, the design that results from those guesses may not be optimal.

In the Bayesian optimal-design approach, the initial guesses do not concentrate on single values. Instead, each parameter is assigned to a prior distribution that may center around the guessed value. The optimality criterion is to minimize the Bayes risk in which the parameters are integrated out in the risk function using the prior distributions.

For generalized linear models, since the exact posterior distributions are often intractable, asymptotic approximations may be used Chaloner and Larntz [42]. The normal approximation to the posterior distribution is commonly used. Several normal approximations are available in Berger. Under easily satisfied assumptions, one approximation to the posterior distribution of μ is to use a multivariate normal distribution

$$N_p\left[\hat{\mu}, I\left(\hat{\mu}, \vartheta\right)^{-1}\right]$$

where $\hat{\mu}$ is the maximum likelihood estimate of μ. A further approximation using the prior distribution of μ as the predictive distribution of $\hat{\mu}$ will be applied to obtain the preposterior expected loss Chaloner and Larntz [41].

The two commonly used optimality criteria for Bayesian design are Bayesian D-optimality and Bayesian A-optimality. The Bayesian D-optimality criterion is to use the expected gain in Shannon information between the posterior and prior distribution as the utility function and choose the design measure ϑ maximizing equation (3) where det stands for the determinant and π^{β} is the prior distribution of β. In the Bayesian A-optimality criterion, which requires that the parameters to be estimated are specified and possibly weighted, we minimize equation (3) where tr(C) is the trace of matrix C and $A(\beta)$ is a symmetric $p \times p$ matrix provided by the specification of what is to be estimated.

Information Matrices for Poisson Models

For a design measure, X, on Y putting p_i weight at k distinct design points x_i, i = 1, · · · k, $\sum n_i = \infty$. In general, the Fisher information matrix $I(\beta, X)$ for the generalized linear regression model can be written as

$$I(\beta, X) = \sum_{i=1}^{k} p_i w_i x_i x_i^T \qquad \ldots (4)$$

where x_i is a p × 1 design vector of the i^{th} design points,

For the one-variable poisson regression model, the Fisher information matrix can be written as

$$I(\beta_0, \beta_1, X) = \begin{pmatrix} \sum p_i w_i & \sum p_i w_i x_i \\ \sum p_i w_i x_i & \sum p_i w_i x_i^2 \end{pmatrix} \qquad \ldots (5)$$

while for the two-variable logistic regression model, the Fisher information matrix can be written as

$$I\left(\beta_0,\beta_1,\beta_2,X\right)=\begin{pmatrix}\sum p_i w_i & \sum p_i w_i x_{1i} & \sum p_i w_i x_{2i}\\ \sum p_i w_i x_{1i} & \sum p_i w_i x_{1i}^2 & \sum p_i w_i x_{1i} x_{2i}\\ \sum p_i w_i x_{2i} & \sum p_i w_i x_{1i} x_{2i} & \sum p_i w_i x_{2i}^2\end{pmatrix} \quad \ldots (6)$$

Bayesian A-Optimal Design for Poisson Regression Models

Under certain model assumptions and design optimality criteria, the X that maximizes the Fisher information about the unknown μ is selected. This is usually done by choosing the X to optimize a certain function of the expected Fisher information matrix,

$$I\left(\lambda,\vartheta\right)=-E_y\left\{\frac{\partial}{\partial\lambda^T}\left[\frac{\partial\log p\left(y/\lambda,x\right)}{\partial\lambda}\right]\right\} \quad \ldots (7)$$

We examine the Bayesian D-optimal design for some poisson models. The Bayesian D-optimality is given by;

$$\varphi_1\left(Y\right)=E^{\pi(\beta)}\left\{\log\left[\det I\left(\beta,Y\right)\right]\right\} \quad \ldots (8)$$

which selects the design measure X maximizing $\phi_1\left(Y\right)$. We also examine the Bayesian A-optimal design for some poisson models. The poisson regression model is very useful in modelling the count data, equation (1) is a generalized linear model with unknown parameters in the information matrix. The Bayesian A-optimality is given by;

$$\varphi_2\left(Y\right)=-E^{\pi(\beta)}\left\{tr\left[A\left(\beta\right)I\left(\beta,Y\right)^{-1}\right]\right\} \quad \ldots (9)$$

which selects the design measure X maximizing $\phi_2\left(Y\right)$. Assuming that the experimenter doesn't have much knowledge about the parameters, a range of uniform and independent prior distributions for the parameters are used to find the Bayesian optimal design points.

Efficiency of the Bayesian D-Optimal Design

The goal of the Bayesian D-optimal design is to find design points at which the determinant

of the Fisher information matrix evaluated at the true parameter values is maximized. The

D-efficiency is defined as the ratio of the determinant of the Fisher information matrix with the chosen design points to that with the true D-optimal design points at the true parameter values, i.e.,

$$\xi_1(Y) = D-eff = \frac{\left|I\left(Y,\beta_{true}\right)\right|}{\left|I\left(Y_{D-opt},\beta_{true}\right)\right|} \qquad ...(10)$$

Efficiency of the Bayesian A-Optimal Design

The goal of the Bayesian A-optimal design is to find design points at which the trace of the Fisher information matrix evaluated at the true parameter values is maximized. The A-efficiency is defined as the ratio of the determinant of the Fisher information matrix with the chosen design points to that with the true A-optimal design points at the true parameter values, i.e.,

$$\xi_2(X) = \frac{tr\left(X,\beta_{true}\right)}{I\left(X_{A-opt,}\beta_{true}\right)} \qquad ...(11)$$

3. Results

Likelihood:

CGPA ~ poisson(xb_CGPA)

Prior:

{CGPA:AGE STUDY INTERNET SLEEP LIVING _cons} ~ normal(0,10000) (1)

----

(1) Parameters are elements of the linear form xb_CGPA.

Bayesian Poisson regression	MCMC iterations	=	12,500
Random-walk Metropolis-Hastings sampling	Burn-in	=	2,500
	MCMC sample size	=	10,000
	Number of obs	=	300
	Acceptance rate	=	.1828

				Efficiency: min	=	.02325
				avg	=	.03809
Log marginal likelihood = -564.18065				max	=	.05763

					Equal-tailed	
VAR	IRR	Std. Dev.	MCSE	Median	[95% Cred. Interval]	
AGE	.9984969	.0082223	.000452	.9982041	.982632	1.01407
STUDY	.9933854	.0077997	.000485	.9930523	.9787901	1.009298
INTERNET	.9957648	.0083724	.00043	.9951907	.9798274	1.013192
SLEEP	.9990856	.0081169	.000338	.9986993	.9831727	1.015592
LIVING	1.006449	.0398522	.002613	1.004643	.9277565	1.090451
_cons	4.013821	1.048096	.046447	3.859159	2.406909	6.408838

The OPTEX Procedure

Table 1

Factor Ranges

Factor	**Low Value**	**High Value**
x1	-1.000000	1.000000
x2	-1.000000	1.000000
x3	-1.000000	1.000000
x4	-1.000000	1.000000
x5	-1.000000	1.000000
x6	-1.000000	1.000000
x7	-1.000000	1.000000
x8	-1.000000	1.000000
x9	-1.000000	1.000000

The OPTEX Procedure

Table 2

Design Number	D-Efficiency	A-Efficiency	G-Efficiency	Average Prediction Standard Error
1	99.8014	99.6019	98.1318	0.3845*
2	99.7854	99.5697	98.0099	0.3846
3	99.7799	99.5598	97.8837	0.3846
4	99.7761	99.5520	98.3159	0.3846
5	99.7729	99.5466	98.4162	0.3846
6	99.7706	99.5408	98.0497	0.3846
7	99.7591	99.5191	98.4355	0.3847
8	99.7546	99.5100	98.0490	0.3847
9	99.7533	99.5060	97.9757	0.3847
10	99.7510	99.5011	97.9052	0.3847

The GLM Procedure

Table 3

General Form of Aliasing Structure
Intercept
x1
x2
x3
x4
x5
x6
x7
x8
x9
x1*x2
x1*x3
x2*x3
x1*x4
x2*x4
x3*x4
x1*x5
x2*x5
x3*x5
x4*x5
x1*x6
x2*x6
x3*x6
x4*x6
x5*x6
x1*x7
x2*x7
x3*x7

General Form of Aliasing Structure

x4*x7
x5*x7
x6*x7
x1*x8
x2*x8
x3*x8
x4*x8
x5*x8
x6*x8
x7*x8
x1*x9
x2*x9
x3*x9
x4*x9
x5*x9
x6*x9
x7*x9
x8*x9

The GLM Procedure

Dependent Variable: y

Table 4

Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	45	5.12359391	0.11385764	1.42	0.0510
Error	254	20.41040795	0.08035594		
Corrected Total	299	25.53400186			

Table 5

R-Square	**Coeff Var**	**Root MSE**	**y Mean**
0.200658	55.79938	0.283471	0.508019

Table 6

SOURCE	**DF**	**TYPE I SS**	**MEAN SQUARE**	**F VALUE**	**PR > F**
X1	1	0.04964580	0.04964580	0.62	0.4326
X2	1	0.07427179	0.07427179	0.92	0.3373
X3	1	0.10984251	0.10984251	1.37	0.2434
X4	1	0.28201230	0.28201230	3.51	0.0622
X5	1	0.12803773	0.12803773	1.59	0.2080
X6	1	0.52375595	0.52375595	6.52	0.0113

X7	1	0.03863451	0.03863451	0.48	0.4887
X8	1	0.00420088	0.00420088	0.05	0.8193
X9	1	0.00235324	0.00235324	0.03	0.8643
X1*X2	1	0.22017450	0.22017450	2.74	0.0991
X1*X3	1	0.43775739	0.43775739	5.45	0.0204
X2*X3	1	0.00247002	0.00247002	0.03	0.8610
X1*X4	1	0.01424717	0.01424717	0.18	0.6741
X2*X4	1	0.00276176	0.00276176	0.03	0.8531
X3*X4	1	0.07215566	0.07215566	0.90	0.3442
X1*X5	1	0.41249580	0.41249580	5.13	0.0243
X2*X5	1	0.02551130	0.02551130	0.32	0.5736
X3*X5	1	0.05006306	0.05006306	0.62	0.4307
X4*X5	1	0.08729260	0.08729260	1.09	0.2983
X1*X6	1	0.00142262	0.00142262	0.02	0.8943
X2*X6	1	0.02615450	0.02615450	0.33	0.5688
X3*X6	1	0.22532900	0.22532900	2.80	0.0953
X4*X6	1	0.00028789	0.00028789	0.00	0.9523
X5*X6	1	0.00060841	0.00060841	0.01	0.9307
X1*X7	1	0.21575110	0.21575110	2.68	0.1025
X2*X7	1	0.12809671	0.12809671	1.59	0.2079
X3*X7	1	0.25766852	0.25766852	3.21	0.0745
X4*X7	1	0.03271566	0.03271566	0.41	0.5240
X5*X7	1	0.00530801	0.00530801	0.07	0.7974
X6*X7	1	0.04981625	0.04981625	0.62	0.4318
X1*X8	1	0.09981079	0.09981079	1.24	0.2661
X2*X8	1	0.00491921	0.00491921	0.06	0.8048
X3*X8	1	0.01369067	0.01369067	0.17	0.6801
X4*X8	1	0.32465854	0.32465854	4.04	0.0455
X5*X8	1	0.08339143	0.08339143	1.04	0.3093
X6*X8	1	0.01360928	0.01360928	0.17	0.6810
X7*X8	1	0.13362799	0.13362799	1.66	0.1984
X1*X9	1	0.09445953	0.09445953	1.18	0.2793
X2*X9	1	0.10357124	0.10357124	1.29	0.2573
X3*X9	1	0.42744545	0.42744545	5.32	0.0219
X4*X9	1	0.00019049	0.00019049	0.00	0.9612
X5*X9	1	0.02656552	0.02656552	0.33	0.5658
X6*X9	1	0.03443808	0.03443808	0.43	0.5133
X7*X9	1	0.17388753	0.17388753	2.16	0.1425
X8*X9	1	0.10848552	0.10848552	1.35	0.2464

SOURCE	DF	TYPE III SS	MEAN SQUARE	F VALUE	PR > F
X1	1	0.02941719	0.02941719	0.37	0.5457
X2	1	0.06687261	0.06687261	0.83	0.3625
X3	1	0.10025676	0.10025676	1.25	0.2651
X4	1	0.31944642	0.31944642	3.98	0.0472
X5	1	0.12426009	0.12426009	1.55	0.2148
X6	1	0.56153743	0.56153743	6.99	0.0087
X7	1	0.03093564	0.03093564	0.38	0.5355
X8	1	0.00850194	0.00850194	0.11	0.7452
X9	1	0.00023349	0.00023349	0.00	0.9571
X1*X2	1	0.21611042	0.21611042	2.69	0.1023
X1*X3	1	0.44456002	0.44456002	5.53	0.0194
X2*X3	1	0.00106261	0.00106261	0.01	0.9085
X1*X4	1	0.00994286	0.00994286	0.12	0.7253
X2*X4	1	0.00559911	0.00559911	0.07	0.7920
X3*X4	1	0.07090010	0.07090010	0.88	0.3485
X1*X5	1	0.47608897	0.47608897	5.92	0.0156
X2*X5	1	0.02819740	0.02819740	0.35	0.5541
X3*X5	1	0.04581569	0.04581569	0.57	0.4509
X4*X5	1	0.09940774	0.09940774	1.24	0.2671
X1*X6	1	0.00017169	0.00017169	0.00	0.9632
X2*X6	1	0.02187133	0.02187133	0.27	0.6023
X3*X6	1	0.22142050	0.22142050	2.76	0.0982
X4*X6	1	0.00069580	0.00069580	0.01	0.9259
X5*X6	1	0.00036603	0.00036603	0.00	0.9462
X1*X7	1	0.21890413	0.21890413	2.72	0.1001
X2*X7	1	0.11879012	0.11879012	1.48	0.2252
X3*X7	1	0.24831177	0.24831177	3.09	0.0800
X4*X7	1	0.04079283	0.04079283	0.51	0.4768
X5*X7	1	0.00650185	0.00650185	0.08	0.7763
X6*X7	1	0.04701798	0.04701798	0.59	0.4450
X1*X8	1	0.08077053	0.08077053	1.01	0.3170
X2*X8	1	0.00483870	0.00483870	0.06	0.8064
X3*X8	1	0.01194486	0.01194486	0.15	0.7002
X4*X8	1	0.31283789	0.31283789	3.89	0.0496
X5*X8	1	0.08353019	0.08353019	1.04	0.3089
X6*X8	1	0.01479496	0.01479496	0.18	0.6682
X7*X8	1	0.13104280	0.13104280	1.63	0.2028
X1*X9	1	0.08971828	0.08971828	1.12	0.2917

X2*X9	1	0.09779981	0.09779981	1.22	0.2710
X3*X9	1	0.43646782	0.43646782	5.43	0.0206
X4*X9	1	0.00013836	0.00013836	0.00	0.9669
X5*X9	1	0.02461186	0.02461186	0.31	0.5805
X6*X9	1	0.03433435	0.03433435	0.43	0.5139
X7*X9	1	0.17722474	0.17722474	2.21	0.1388
X8*X9	1	0.10848552	0.10848552	1.35	0.2464

PARAMETER	ESTIMATE	STANDARD ERROR	T VALUE	PR > |T|	EXPECTED VALUE
INTERCEPT	0.5085373644	0.01638068	31.04	<.0001	Intercept
X1	-.0099319424	0.01641507	-0.61	0.5457	x1
X2	0.0149605253	0.01639954	0.91	0.3625	x2
X3	-.0183248420	0.01640561	-1.12	0.2651	x3
X4	0.0327061012	0.01640358	1.99	0.0472	x4
X5	-.0203946538	0.01640059	-1.24	0.2148	x5
X6	-.0433388470	0.01639445	-2.64	0.0087	x6
X7	-.0101894905	0.01642222	-0.62	0.5355	x7
X8	-.0053342915	0.01639936	-0.33	0.7452	x8
X9	0.0008839163	0.01639772	0.05	0.9571	x9
X1*X2	0.0268932339	0.01639888	1.64	0.1023	x1*x2
X1*X3	-.0386158002	0.01641757	-2.35	0.0194	x1*x3
X2*X3	0.0018850493	0.01639247	0.11	0.9085	x2*x3
X1*X4	0.0057699477	0.01640307	0.35	0.7253	x1*x4
X2*X4	0.0043262747	0.01638942	0.26	0.7920	x2*x4
X3*X4	-.0154025037	0.01639747	-0.94	0.3485	x3*x4
X1*X5	-.0399273684	0.01640347	-2.43	0.0156	x1*x5
X2*X5	0.0097202282	0.01640895	0.59	0.5541	x2*x5
X3*X5	0.0123909696	0.01640995	0.76	0.4509	x3*x5
X4*X5	0.0182457501	0.01640441	1.11	0.2671	x4*x5
X1*X6	-.0007582876	0.01640484	-0.05	0.9632	x1*x6
X2*X6	-.0085570414	0.01640193	-0.52	0.6023	x2*x6
X3*X6	-.0272254610	0.01640119	-1.66	0.0982	x3*x6
X4*X6	0.0015242913	0.01638073	0.09	0.9259	x4*x6
X5*X6	-.0011068709	0.01640025	-0.07	0.9462	x5*x6
X1*X7	-.0270963784	0.01641698	-1.65	0.1001	x1*x7
X2*X7	0.0199420562	0.01640170	1.22	0.2252	x2*x7
X3*X7	0.0288233528	0.01639664	1.76	0.0800	x3*x7
X4*X7	0.0116915061	0.01640920	0.71	0.4768	x4*x7
X5*X7	0.0046675299	0.01640883	0.28	0.7763	x5*x7

X6*X7	0.0125467284	0.01640240	0.76	0.4450	x6*x7
X1*X8	0.0164361859	0.01639395	1.00	0.3170	x1*x8
X2*X8	0.0040221778	0.01639101	0.25	0.8064	x2*x8
X3*X8	0.0063219655	0.01639723	0.39	0.7002	x3*x8
X4*X8	0.0323496185	0.01639527	1.97	0.0496	x4*x8
X5*X8	0.0167184006	0.01639766	1.02	0.3089	x5*x8
X6*X8	0.0070330975	0.01639074	0.43	0.6682	x6*x8
X7*X8	0.0209633936	0.01641587	1.28	0.2028	x7*x8
X1*X9	-.0173379043	0.01640836	-1.06	0.2917	x1*x9
X2*X9	-.0180974000	0.01640424	-1.10	0.2710	x2*x9
X3*X9	-.0382190606	0.01639883	-2.33	0.0206	x3*x9
X4*X9	-.0006800792	0.01638926	-0.04	0.9669	x4*x9
X5*X9	0.0090797100	0.01640623	0.55	0.5805	x5*x9
X6*X9	-.0107198748	0.01639964	-0.65	0.5139	x6*x9
X7*X9	0.0243730487	0.01641182	1.49	0.1388	x7*x9
X8*X9	-.0190628170	0.01640628	-1.16	0.2464	x8*x9

Bayesian Logistic Regression for a Single Variable

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits	
Intercept	1	1.2589	0.2121	0.8432	1.6746
AGE	1	-0.0011	0.0081	-0.0170	0.0147
Scale	0	1.0000	0.0000	1.0000	1.0000

Initial Values of the Chain

Chain	**Seed**	**Intercept**	**AGE**
1	1	1.258917	-0.00114

Fit Statistics

DIC (smaller is better)	1030.035
pD (effective number of parameters)	2.006

Posterior Intervals

Parameter	**Alpha**	**Equal-Tail Interval**		**HPD Interval**	
Intercept	0.050	0.8352	1.6753	0.8490	1.6847
AGE	0.050	-0.0170	0.0149	-0.0167	0.0150

Geweke Diagnostics

Parameter	z	Pr > |z|
Intercept	-0.7591	0.4478
AGE	0.5152	0.6064

Effective Sample Sizes

Parameter	ESS	Autocorrelation Time	Efficiency
Intercept	8199.2	1.2196	0.8199
AGE	8221.1	1.2164	0.8221

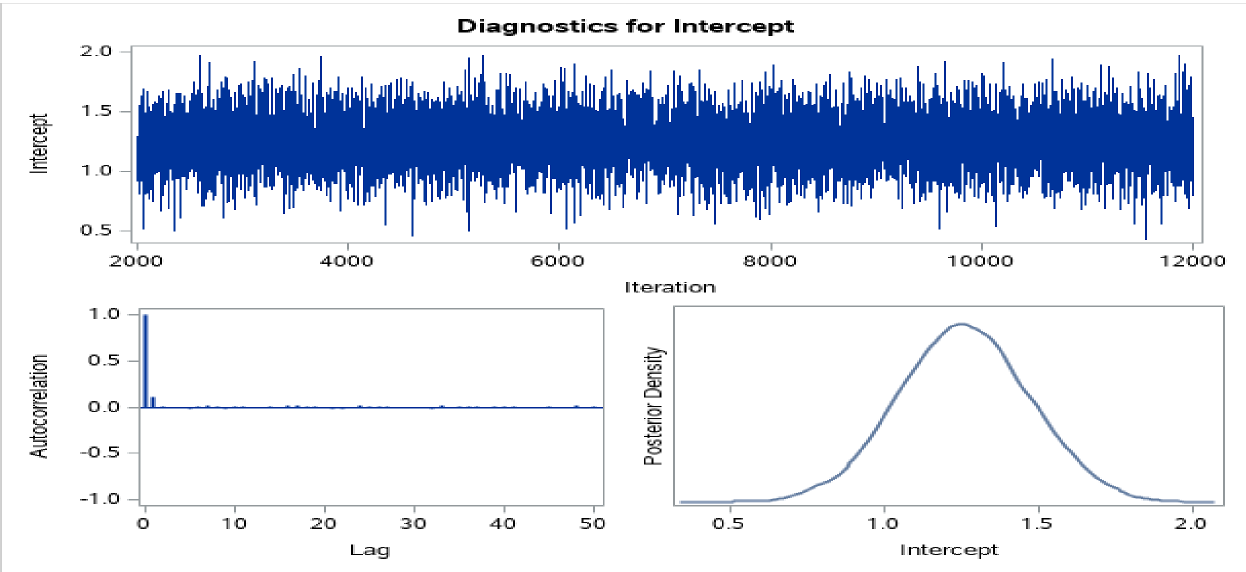

Figure 1: In the panel of the diagnostic plots above, the first graph shows sparsely good spikes for the posterior distribution of the intercept. Autocorrelations are high in the first ten lags but low towards the end, and the posterior density is approximately normal and a bit smooth.

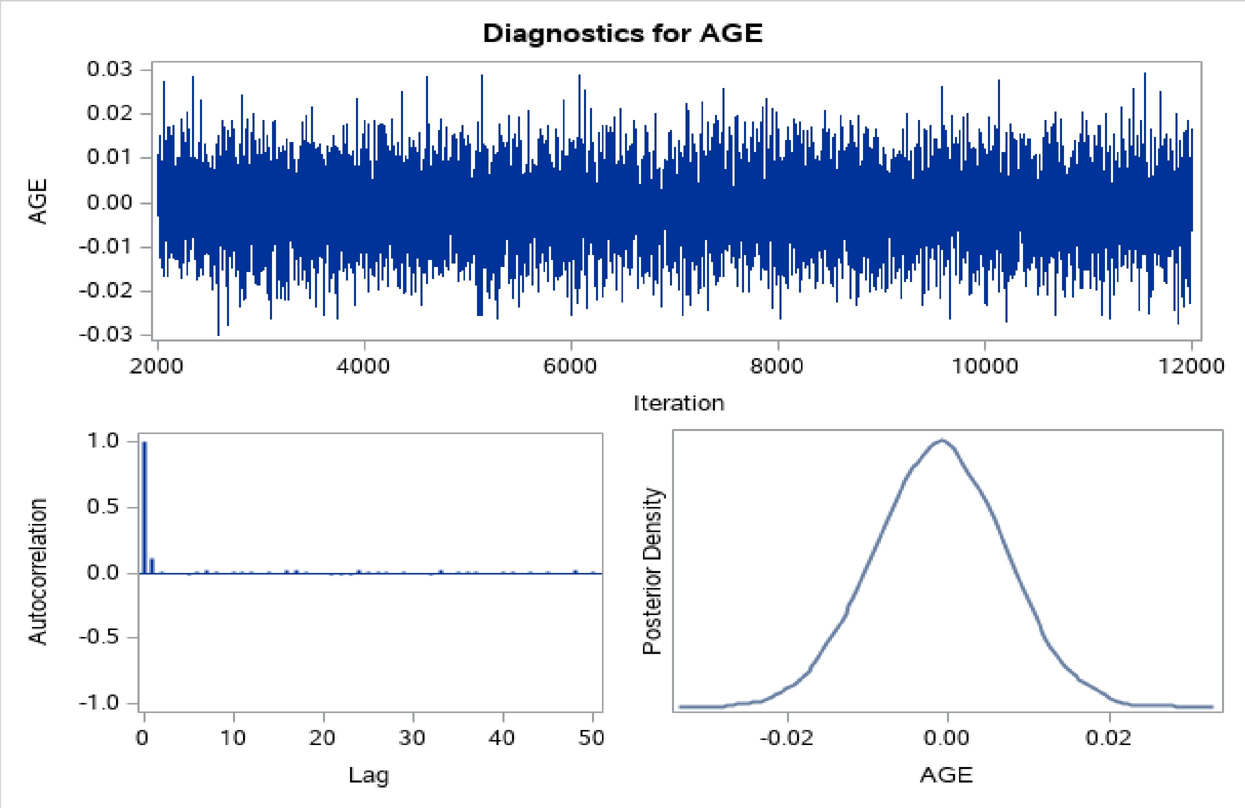

Figure 2: In the panel of the diagnostic plots above, the first graph shows sparsely good spikes for the posterior distribution of the Age. Autocorrelations are slightly high in the first ten lags but low towards the end, and the posterior density is approximately normal and a bit smooth.

Bayesian Logistic Regression for Two Variables

Table 6

ANALYSIS OF MAXIMUM LIKELIHOOD PARAMETER ESTIMATES

PARAMETER	DF	Estimate	Standard Error	Wald 95% Confidence Limits	
INTERCEPT	1	1.3288	0.2230	0.8916	1.7659
AGE	1	-0.0013	0.0081	-0.0172	0.0145
STUDY	1	-0.0077	0.0077	-0.0227	0.0074
SCALE	0	1.0000	0.0000	1.0000	1.0000

Initial Values of the Chain

Chain	Seed	Intercept	AGE	STUDY
1	1	1.328775	-0.00134	-0.00765

Fit Statistics

DIC (smaller is better)	1031.042
pD (effective number of parameters)	3.003

Bayesian Analysis

POSTERIOR SUMMARIES

PARAMETER	N	Mean	Standard Deviation	Percentiles 25%	50%	75%
INTERCEPT	10000	1.3241	0.2257	1.1716	1.3220	1.4750
AGE	10000	-0.00122	0.00815	-0.00669	-0.00124	0.00430
STUDY	10000	-0.00758	0.00764	-0.0127	-0.00759	-0.00251

POSTERIOR INTERVALS

PARAMETER	Alpha	Equal-Tail Interval		HPD Interval	
INTERCEPT	0.050	0.8826	1.7677	0.8962	1.7796
AGE	0.050	-0.0175	0.0147	-0.0171	0.0148
STUDY	0.050	-0.0226	0.00738	-0.0235	0.00644

POSTERIOR CORRELATION MATRIX

PARAMETER	Intercept	AGE	STUDY
INTERCEPT	1.000	-0.949	-0.318
AGE	-0.949	1.000	0.037
STUDY	-0.318	0.037	1.000

POSTERIOR AUTOCORRELATIONS

PARAMETER	Lag 1	Lag 5	Lag 10	Lag 50
INTERCEPT	0.1497	-0.0047	-0.0006	0.0107
AGE	0.1491	-0.0022	-0.0056	0.0187
STUDY	0.1393	-0.0111	-0.0101	0.0028

EFFECTIVE SAMPLE SIZES

PARAMETER	ESS	Autocorrelation Time	Efficiency
INTERCEPT	7431.8	1.3456	0.7432
AGE	7485.3	1.3359	0.7485
STUDY	7297.1	1.3704	0.7297

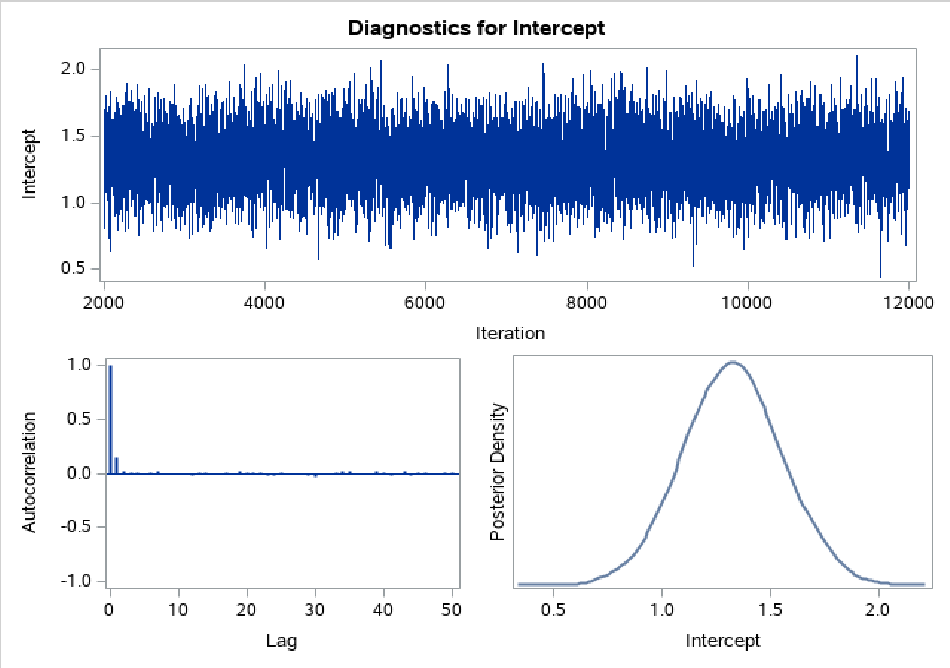

Figure 3: In the panel of the diagnostic plots above, the first graph shows sparsely good spikes for the posterior distribution of the intercept. Autocorrelations are high in the first ten lags but low towards the end, and the posterior density is approximately normal and a bit smooth.

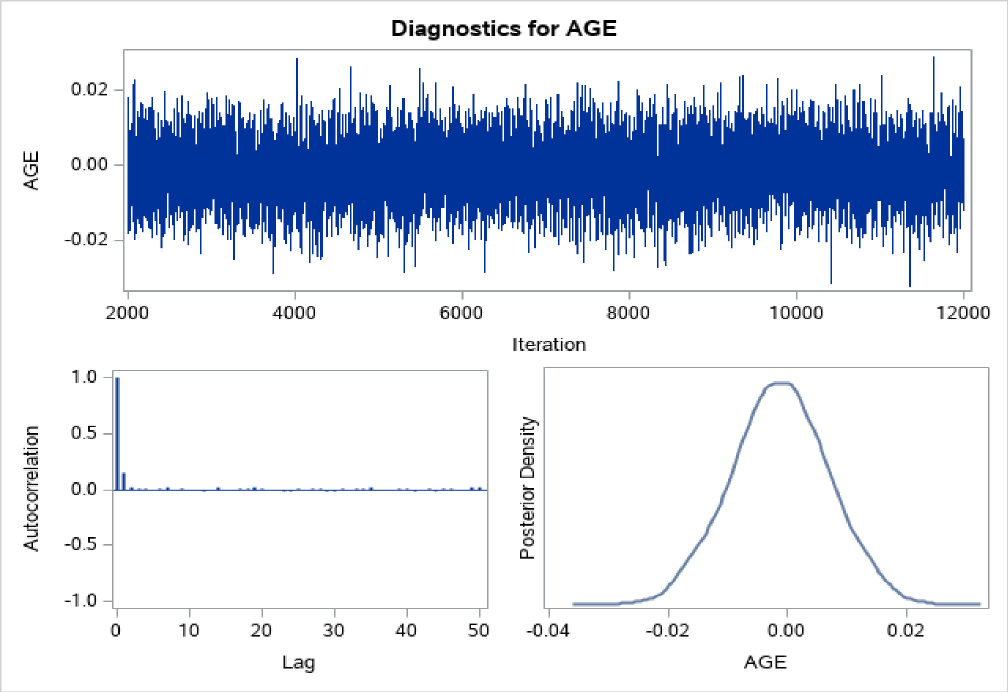

Figure 4: In the panel of the diagnostic plots above, the first graph shows sparsely good spikes for the posterior distribution of the Age. Autocorrelations are slightly high in the first ten lags but low towards the end, and the posterior density is approximately normal and a bit smooth.

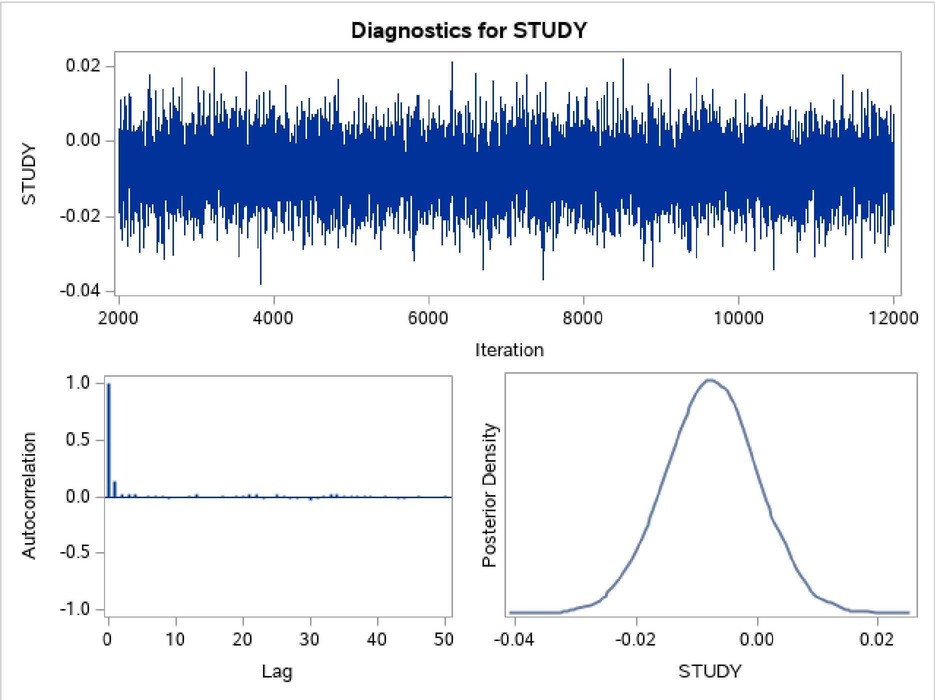

Figure 5: In the panel of the diagnostic plots above, the first graph shows sparsely good spikes for the posterior distribution of the Study. Autocorrelations are slightly high in the first ten lags but low towards the end, and the posterior density is approximately normal and a bit smooth.

4. Conclusion

In this paper, a Bayesian optimal design framework is implemented for D- & A-Optimality using SAS. Bayesian D- & A-Optimality criteria is derived based on expected Shannon information gain on the optimum point. To evaluate the proposed criteria, an algorithm to evaluate the analytically intractable design criterion is used. Bayesian Poisson regression has the benefit that it gives us a posterior distribution rather than a single point estimate. The Bayes Poisson showed that the best design was found in first try out of ten with a D- and A-efficiency of 99.8014 and 99.6019 respectively and positive optimality. A survey was conducted on 300 students to determine the factors that influence academic performance, questionnaires were distributed and their files were assessed from their respective departments to ascertained their actual cumulative grade point average. The data were subjected to statistical analysis, the results showed that Age, mode of stay and study time positively affect the incidence of cumulative grade point average of students, the posterior density is approximately normal and a bit smooth on all the parameters under consideration.

References

[1] Ali, H. U. Lasisi, K. E. and Nwaosu, S. C. (2017). Comparing the Performance of Bayesian and Frequentist Analysis Methods of Irregular Fractional Factorials Using Design Based Optimality and Efficiency Criteria. *IOSR Journal of Mathematics*, 13, 00-00.

[2] Atkinson, A. C., A. N. Donev, and R. D. Tobias (2007). *Optimum Experimental Design, with SAS* (2nd ed.). Oxford: Oxford University Press.

[3] Atkinson, A. C. and D. C. Woods (2015). Designs for generalized linear models. In A. M. Dean, M. D. Morris, J. Stufken, and D. R. Bingham (Eds.), *Handbook of Design and Analysis of Experiments*. Boca Raton: Chapman & Hall/CRC.

[4] Chaloner, K., and Larntz, K. (1989). Optimal Bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference.* 21, 191–208.

[5] Chaloner, K. and I. Verdinelli (1995). Bayesian experimental design: a review. *Statistical Science 10*, 273–304.

[6] Michel, K. (2020). Introduction to Bayesian Logistic Regression. *Datascience* 5, 210-221.

[7] Ryan, E. G., C. C. Drovandi, J. M. McGree, and A. N. Pettitt (2016). A review of modern computational algorithms for Bayesian optimal design. *International Statistical Review 84*, 128–154.

[8] Qiu, J. (2014). Finding optimal experimental designs for models in biomedical studies via particle swarm optimization. PhD thesis, UCLA (2014). https://escholarship.org/uc/item/1cj4b854