World Journal of Innovative Research (WJIR)
ISSN: 2454-8236, Volume-12, Issue-1, January 2022 Pages 13-25

Bayesian D & A Optimal Designs for Logistic Regression Model with Biomedical Application

Ali H., Nwaosu S. C., Lasisi K. E., Abdulkadir A.

***Abstract*— Bayesian optimal designs for binary responses analyzed with logistic regression describing a linear health effect were considered. To overcome the problem of dependence of Bayesian designs on the choice of prior distributions, Bayesian D & A-optimal designs were proposed for logistic regression model. The results show that the optimal number of time points depends on the subject-to-measurement cost ratio and increases with the cost ratio. Furthermore, Bayesian D & A-optimal designs are highly efficient and robust under changes in priors. When implementing the efficiencies of designs with the Bayesian D- & A-optimal designs in modelling chronic heart disease, age and body mass index, it was found that age and BMI are significant in patients heart disease.**

***Index Terms*— Bayesian Logistic Regression, D-Optimality, A-Optimality, Chronic Heart Disease.**

I. INTRODUCTION

Experimental Design is an a priori concept, taking place before data has been collected, and hence the Bayesian paradigm is a particular 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 alsoreduce reliance on unrealistic assumptions and asymptoticresults that may be inappropriate for small tomedium-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. 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 et al, (2007).

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 notevolved far from the methods reviewed by Chalonerand Verdinelli (1995). Development and applicationof methods for Bayesian design have lagged behindthe progress made in inference and modelling dueto the additional complexity introduced by the needto 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 et al. (2016).

Design of Experiments for binary responses are very important in biological and clinical trials. Discussion of the non-Bayesian design for logistic regression models can be found in Finney (1999). 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. Chaloner and Larntz (1989) examined the Bayesian optimal design for the one-variable logistic regression model using the Nelder-Mead algorithm. 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 multi-variable logistic regression models.

Ali H., Department of Mathematics, University of Jos, Nigeria
Nwaosu S. C., Department of Statistics, Joseph Sawuaan Tarka University of Agriculture, Makurdi, Nigeria
Lasisi K. E., Department of Mathematical Sciences, Abubakar Tafawa Balewa University,P. M. B. 0248, Bauchi, Nigeria.
Abdulkadir A., Department of Mathematical Sciences, Abubakar Tafawa Balewa University,, P. M. B. 0248, Bauchi, Nigeria

II. MATERIALS AND METHODS

We examine the Bayesian D-optimal design for some logistic models. The logistic regression model is very useful in modelling the binary responses, in a generalized linear model with unknown parameters in the information matrix. The Bayesian D-optimality is given by;

$$\varphi_1(X) = E^{\pi(\beta)}\left\{\log\left[\det I(\beta, X)\right]\right\} \quad \ldots (1)$$

which selects the design measure X maximizing $\phi_1(X)$. Assuming that the experimenter does not know much information about the parameters, a range of uniform and independent prior distributions for the parameters are used to find the Bayesian optimal design points.

We also examine the Bayesian A-optimal design for some logistic models. The logistic regression model which is very useful in modelling the binary responses, is a generalized linear model with unknown parameters in the information matrix. The Bayesian A-optimality is given by;

$$\varphi_2(X) = -E^{\pi(\beta)}\left\{tr\left[A(\beta)I(\beta, X)^{-1}\right]\right\} \quad \ldots (2)$$

which selects the design measure X maximizing $\phi_2(X)$. 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.

The logistic regression model can be written as

$$y_{ij} \sim Logistic(n_i),$$

$$where$$

$$n_i = \frac{1}{1+\exp(-x_i^T \beta)} \quad \ldots (3)$$

We define y_{ij} in (3) to be response for the j^{th} replicate of the i^{th} design point and assume it follows a logistic distribution with n_i as the mean; x_i is the regressor vector at the i^{th} point and β is the parameter vector. For the one-variable model and two-variable additive model,

$$n_i = \frac{1}{1+\exp(-\beta_0 - \beta_1 x_i)} \quad \ldots (4)$$

$$n_i = \frac{1}{1+\exp(-\beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i})} \quad \ldots (5)$$

Denote by p_i the proportion of whole sample size at i^{th} design point x_i, hence $\sum n_i = 1.$

Also write $w_i = n_i(1-n_i)$ for $i = 1, 2, ..., k.$ The Fisher information matrix $I(\beta, X)$ for the logistic regression model can be written as

$$I(\beta, X) = \sum_{i=1}^{k} p_i w_i x_i x_i^T \quad \ldots (6)$$

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

$$D-eff = \frac{|I(X, \beta_{true})|}{|I(X_{D-opt}, \beta_{true})|} \quad \ldots (7)$$

$$\xi_1(X) =$$

INFORMATION MATRICES FOR LOGISTIC MODELS

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

$$I(\beta, X) = \sum_{i=1}^{k} p_i w_i x_i x_i^T \quad \ldots (8)$$

where x_i is a p × 1 design vector of the i^{th} design points, and $w_i = n_i(1-n_i)$ for logistic regression models.

A general logistic regression model is given in (3) and its associated two different models, namely, the one-variable model and two-variable model, are given in equations (4)–(5), respectively.

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

$$I(\beta_0, \beta_1, X) = \begin{pmatrix} \sum p_i w_i & \sum p_i w_i x_i \\ \sum p_i w_i x_i & \sum p_i w_i x_i^2 \end{pmatrix} \quad \ldots (9)$$

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

$$I(\beta_0, \beta_1, \beta_2, X) = \begin{pmatrix} \sum p_i w_i & \sum p_i w_i x_{1i} & \sum p_i w_i x_{2i} \\ \sum p_i w_i x_{1i} & \sum p_i w_i x_{1i}^2 & \sum p_i w_i x_{1i} x_{2i} \\ \sum p_i w_i x_{2i} & \sum p_i w_i x_{1i} x_{2i} & \sum p_i w_i x_{2i}^2 \end{pmatrix} \quad \ldots (10)$$

Consider the broad class of models for which the response, *Y*, follows a logistic distribution with the expectation $P(\alpha + \beta_i x)$, where *P* is a cumulative distribution function. For the estimation of α and β_i, the exact optimal design problem is to choose *k* distinct $x_1 \ldots, x_k$ and $n_i$observations on each of x_i *with* respect to some optimality criterion for fixed *n*. Here $\sum_{i=1}^{k} n_i = n.$ Since this is a difficult and often intractable optimization problem, the corresponding approximate design, in which n_i/n is replaced by ξ_i, is considered. Thus, a design can be denoted by $d = \{(x_i, \xi_i), i = 1, ..., k\}$, , where ξ_i >0 and $\sum_{i=1}^{k} \xi_i = 1.$ We shall denote the entire class of all such designs by *D*.

It is well known that the information matrix for a given design *d* is

$$I_d(\alpha,\beta_i)=\begin{pmatrix}\sum_{i=1}^{k}\xi_i\psi(c_i) & \sum_{i=1}^{k}\xi_i x_i\psi(c_i)\\ \sum_{i=1}^{k}\xi_i x_i\psi(c_i) & \sum_{i=1}^{k}\xi_i x_i^2\psi(c_i)\end{pmatrix}$$

Here $c_i=\alpha+\beta_i x_i$ and $\psi(c_i)=\left\{\left[P^1(c_i)\right]^2/P(c_i)\left[1-P(c_i)\right]\right\}$. We shall assume that *P* satisfies the following condition.

Conditions (i): the density function *P'* is symmetric about zero; $\psi(0)$ >0 and $\lim_{c\to\infty}\psi(c)=0$; when *c* >0, $\psi(c)$ >0, $\left(\psi^{-1/2}(c)\right)^1$ >0, and $\left(\psi^{-1/2}(c)\right)^{11}$ >0.

Condition (i) is not demanding. In fact, commonly used generalized linear models for binary response, such as logistic, probit, and Poisson models, satisfy condition (i).

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.,

$$\xi_2(X)=\frac{tr(X,\beta_{true})}{I(X_{A-opt},\beta_{true})} \quad \ldots (11)$$

INFORMATION MATRICES FOR LOGISTIC MODELS

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

$$I(\beta,X)=\sum_{i=1}^{k}p_i w_i x_i x_i^T \quad \ldots (12)$$

where x_i is a p × 1 design vector of the i^{th} design points, and $w_i=n_i(1-n_i)$ for logistic regression models.

A general logistic regression model is given in (3) and its associated two different models, namely, the one-variable model and two-variable model, are given in equations (4)–(5), respectively.

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

$$I(\beta_0,\beta_1,X)=\begin{pmatrix}\sum p_i w_i & \sum p_i w_i x_i\\ \sum p_i w_i x_i & \sum p_i w_i x_i^2\end{pmatrix} \quad \ldots (13)$$

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

$$I(\beta_0,\beta_1,\beta_2,X)=\begin{pmatrix}\sum p_i w_i & \sum p_i w_i x_{1i} & \sum p_i w_i x_{2i}\\ \sum p_i w_i x_{1i} & \sum p_i w_i x_{1i}^2 & \sum p_i w_i x_{1i}x_{2i}\\ \sum p_i w_i x_{2i} & \sum p_i w_i x_{1i}x_{2i} & \sum p_i w_i x_{2i}^2\end{pmatrix} \quad \ldots (14)$$

Let $f(x_i)\sim Bernoulli(n_i)$ be the response from the experiment with variable settings xi and

$$\log\left(\frac{n_i}{1-n_i}\right)=\beta_0+\sum_{j=1}^{2}\beta_j x_{ij}+\sum_{j=1}^{2}\sum_{k\geq j}^{2}\beta_{jk}x_{ij}x_{ik}, \quad \ldots (15)$$

Where $\beta_0,\beta_1,\beta_2,\beta_{11},\beta_{12},\beta_{22}$ are unknown parameters to be estimated. Here, $\mu(x_i)=n_i$ and the variance is given by $Var\{\mu(x_i)\}=n_i[1-n_i]$ with $\varphi=1$. To show some Bayesian design concepts, Atkinson and Woods (2015) assumed the following independent prior distributions for the parameters:

$$\beta_1,\beta_2\sim U(2,6),\ \beta_0,\beta_3,\beta_{jk}\sim U(-2,2), \quad \text{for j, k = 1, 2, 3.} \quad \ldots (16)$$

THE HOSMER & LEMESHOW GOODNESS-OF-FIT TEST

The steps of constructing the test:

1. For a given logistic regression model, compute the resulting estimated probabilities for all observations in the model. The steps of constructing the test:

1. For a given logistic regression model, compute the resulting estimated probabilities for all observations in the model

$$\hat{n}_i=\frac{1}{1+\exp(-\hat{\beta}_0-\hat{\beta}_1X_{i1}-\ldots-\hat{\beta}_kX_{ik})}$$

2. Sort the data in increasing order byn_i and create g groups (if possible, g=10 approximate deciles)

3. Compute the total observed number of cases and the total expected number of cases

O_j - observed number of cases in decile j

$E_j=\sum_{i\in Groupj}\hat{n}_i$ - expected number of cases in decile j

4. Compute the test statistic

$$\hat{C}=\sum_{i=1}^{g}\frac{\left(O_j-E_j\right)^2}{E_j\left(1-\frac{E_j}{n_j}\right)}\sim\chi^2_{g-2}$$

under H_0:"good fit"

Large values of $\hat{C}$ (and small p-values) indicate a lack of fit.

Small values of $\hat{C}$ (and large p-values) indicate a good model fit.

III. APPLICATIONS

Study population

All patients that reported at the clinic, Jos University Teaching Hospital (JUTH), Plateau State, Nigeria, that met the inclusion criteria who gave consent were enrolled for the study between 10th June 2021 and 10th December, 2021.

Preparation for data collection

The researchers had audience with the patients on the appointment date and discussed the study procedure, process, import and expected date for commencement/conclusion of the study. Clarifications on any grey areas were sought.

Ethical consideration/approval

Ethical clearance was obtained from the Institutional Health Research Ethical Committee of the Jos University Teaching Hospital. Participants' anonymity and confidentiality were maintained in accordance with the Helsinki Declaration.

Data collection instrument

Weights of all the participants were measured using a brand new calibrated digital bath room weighing scale while heights summarize were measured in meters on a standardized calibrated wall which could be substituted with a standard measuring tape in dire emergencies/the critically ill. All

was without shoes, minimal clothing for weight and less of hair inclusion for height.

Data collection/procedure

All consecutive participants were enrolled for the study at the Clinic Hall. Data was generated from their biodata, measured heights to the nearest 0.01 m which we used a standardized calibrated wall while their weights to the nearest 0.01 kg using a brand-new calibrated bathroom digital weighing scale of 120 kg capacity model number BR 9011, made in China. Measurements were recorded without shoes with minimal clothing during weighing while in measuring heights we adopted measures that involved less hair inclusion. Initials were used to conceal identity.

The result was subjected to statistical analysis.

Variable	Obs	Mean	Std. Dev.	Min	Max
AGE	200	24.62	4.339326	19	34
WEIGHT	200	50.36345	11.21777	29.6	94.9
HEIGHT	200	1.56654	.1199724	1.314	1.87
BMI	200	20.49175	3.990604	13.97753	37.29593
CHD	200	.405	.4921239	0	1

Bayesian Logistic Regression for a Single Variable

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits	
Intercept	1	-1.0725	0.8366	-2.7122	0.5672
Age	1	0.0279	0.0332	-0.0372	0.0929
Scale	1	0.4900	0.0245	0.4443	0.5405

Note: The scale parameter was estimated by maximum likelihood.

Bayesian Analysis

Independent Normal Prior for Regression Coefficients

Parameter	Mean	Precision
Intercept	0	1E-6
Age	0	1E-6

Algorithm converged.

Independent Prior Distributions for Model Parameters

Parameter	Prior Distribution
Dispersion	Proper

Initial Values of the Chain

Chain	Seed	Intercept	Age	Dispersion
1	687853127	-1.07245	0.027861	0.237748

Fit Statistics

DIC (smaller is better)	288.540
pD (effective number of parameters)	3.111

Bayesian Analysis

Posterior Summaries

Parameter	N	Mean	Standard Deviation	Percentiles 25%	50%	75%
Intercept	10000	-1.0942	0.9005	-1.6925	-1.0900	-0.5043
Age	10000	0.0283	0.0359	0.00434	0.0284	0.0519
Dispersion	10000	0.2454	0.0247	0.2280	0.2442	0.2610

Posterior Intervals

Parameter	Alpha	Equal-Tail Interval		HPD Interval	
Intercept	0.050	-2.9028	0.7463	-2.9797	0.5670
Age	0.050	-0.0446	0.0992	-0.0402	0.1017
Dispersion	0.050	0.2016	0.2978	0.1992	0.2942

Posterior Correlation Matrix

Parameter	Intercept	Age	Dispersion
Intercept	1.000	-0.986	0.008
Age	-0.986	1.000	-0.010
Dispersion	0.008	-0.010	1.000

Bayesian Analysis

Posterior Autocorrelations

Parameter	Lag 1	Lag 5	Lag 10	Lag 50
Intercept	0.7410	0.1351	0.0326	0.0385
Age	0.7402	0.1310	0.0291	0.0422
Dispersion	0.0079	0.0028	0.0041	-0.0034

Geweke Diagnostics

Parameter	z	Pr > |z|
Intercept	-0.3658	0.7145
Age	0.5878	0.5567
Dispersion	3.0455	0.0023

Effective Sample Sizes

Parameter	ESS	Autocorrelation Time	Efficiency
Intercept	1380.0	7.2464	0.1380
Age	1404.3	7.1211	0.1404
Dispersion	10000.0	1.0000	1.0000

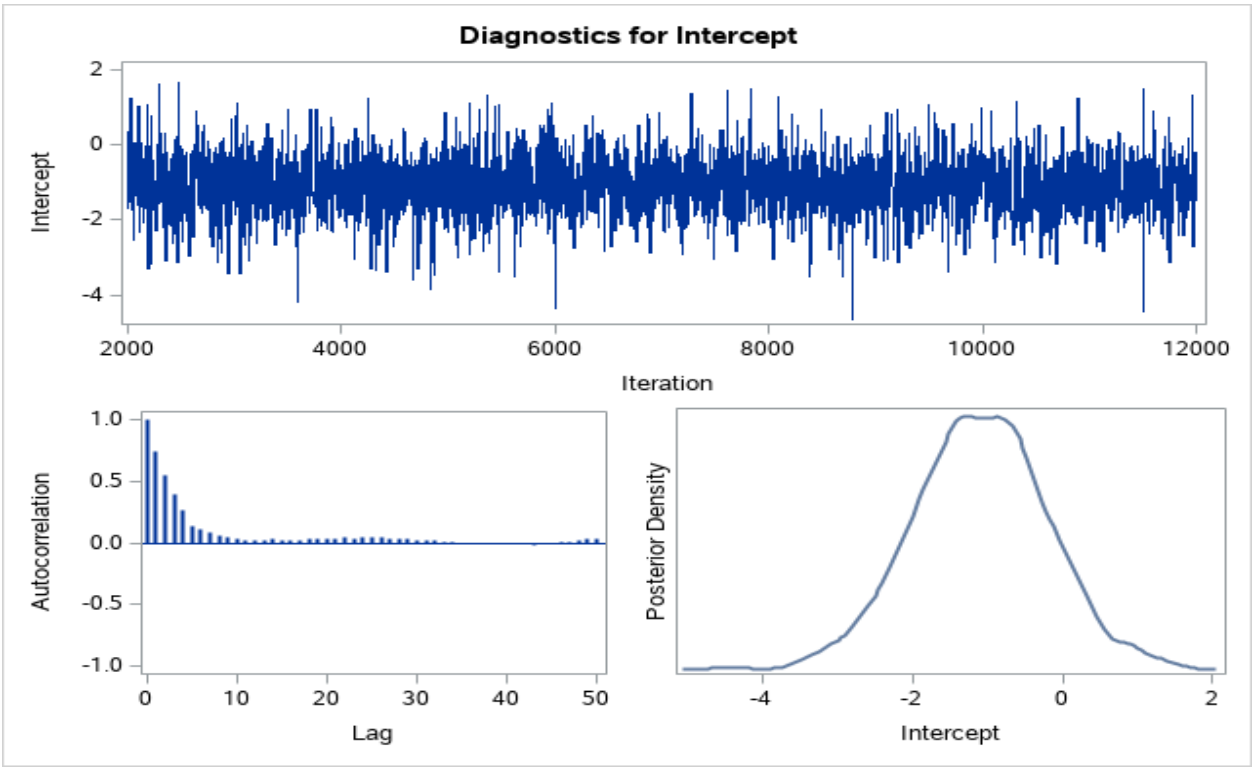

Figure 1. In the panel of the diagnostic plots above, the first graph shows sparsely good spikes for the posterior distribution of the intercept. Autocorrelations are high in the first ten lags but low towards the end, and the posterior density is is approximately normal and a bit smooth.

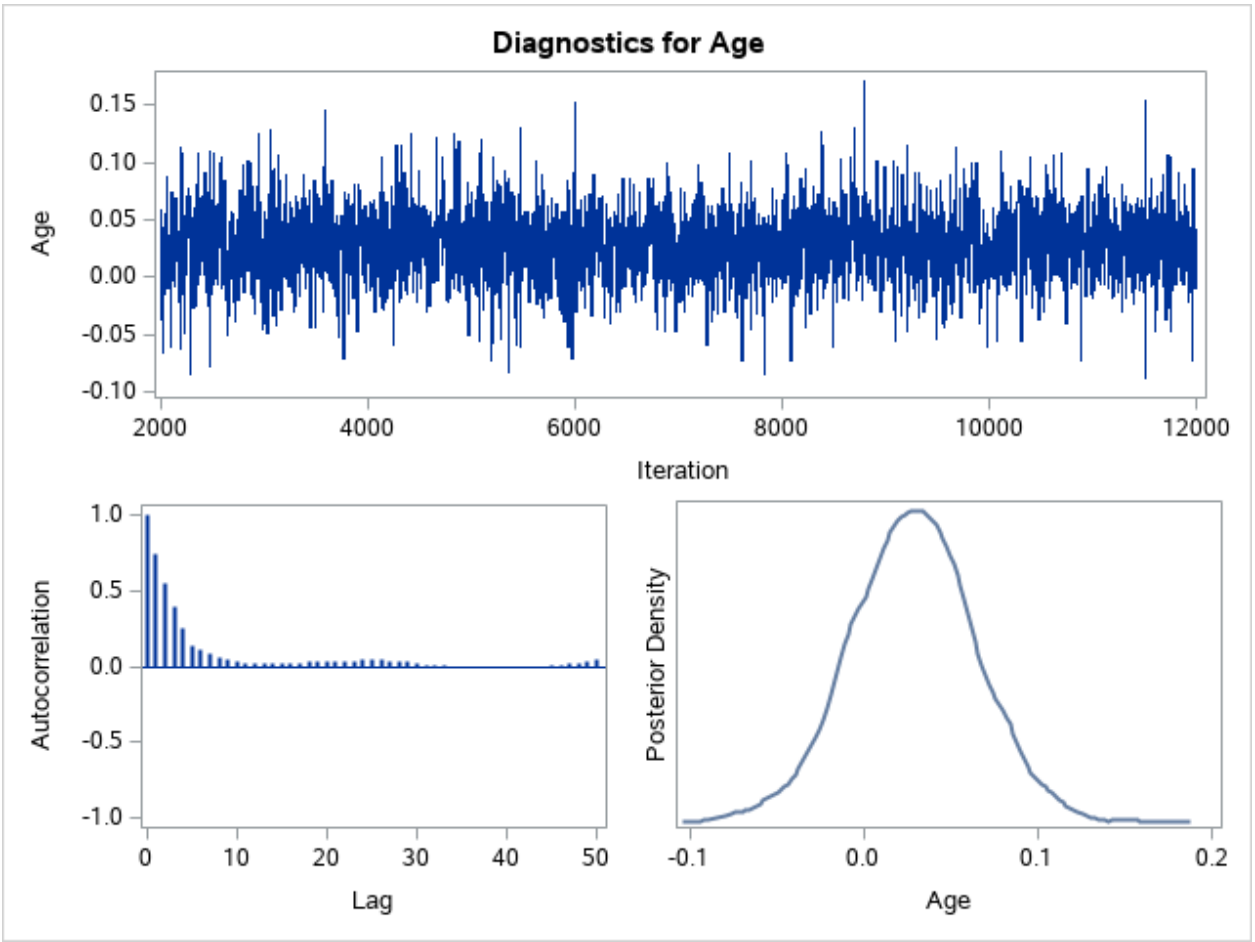

Figure 2: In the panel of the diagnostic plots above, the first graph shows sparsely good spikes for the posterior distribution of the Age. Autocorrelations are slightly high in the first ten lags but low towards the end, and the posterior density is is approximately normal and a bit smooth.

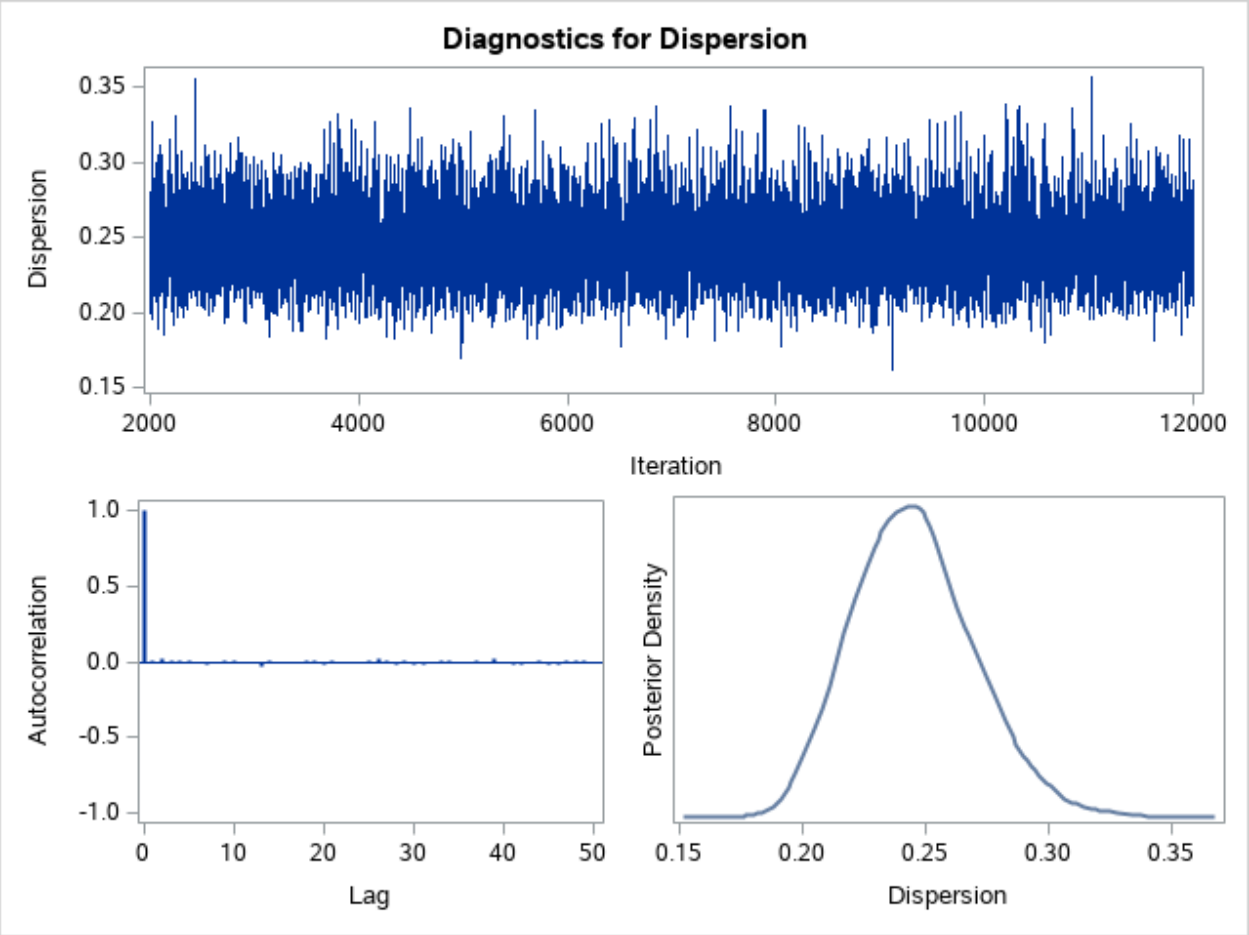

Figure 3: In the panel of the diagnostic plots above, the first graph shows densely good spikes for the posterior distribution of the dispersion. Autocorrelations ishigh in the initial lag but low from lag one to the end, and the posterior density is is approximately normal and bit smooth.

Bayesian Logistic Regression for two Variables

Bayesian Analysis

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits	
Intercept	1	-1.1107	1.1328	-3.3310	1.1096
Age	1	0.0279	0.0332	-0.0371	0.0930
BMI	1	0.0018	0.0359	-0.0686	0.0721
Scale	1	0.4900	0.0245	0.4443	0.5405

Note:The scale parameter was estimated by maximum likelihood.

Bayesian Analysis

Independent Normal Prior for Regression Coefficients

Parameter	Mean	Precision
Intercept	0	1E-6
Age	0	1E-6
BMI	0	1E-6

Algorithm converged.

Independent Prior Distributions for Model Parameters

Parameter	Prior Distribution
Dispersion	Improper

Initial Values of the Chain

Chain	Seed	Intercept	Age	BMI	Dispersion
1	1254751374	-1.11058	0.027914	0.001792	0.237744

Fit Statistics

DIC (smaller is better)	290.663
pD (effective number of parameters)	4.150

Bayesian Analysis

Posterior Summaries

Parameter	N	Mean	Standard Deviation	Percentiles 25%	50%	75%
Intercept	10000	-1.0138	1.2298	-1.8206	-1.0334	-0.2472
Age	10000	0.0266	0.0359	0.00259	0.0263	0.0497
BMI	10000	-0.00230	0.0401	-0.0282	-0.00202	0.0240
Dispersion	10000	0.2465	0.0251	0.2291	0.2448	0.2626

Posterior Intervals

Parameter	Alpha	Equal-Tail Interval		HPD Interval	
Intercept	0.050	-3.5154	1.4830	-3.2465	1.6900
Age	0.050	-0.0455	0.0984	-0.0395	0.1019
BMI	0.050	-0.0833	0.0747	-0.0818	0.0751
Dispersion	0.050	0.2016	0.3005	0.1980	0.2958

Posterior Correlation Matrix

Parameter	Intercept	Age	BMI	Dispersion
Intercept	1.000	-0.736	-0.686	0.006
Age	-0.736	1.000	0.028	0.007
BMI	-0.686	0.028	1.000	-0.019
Dispersion	0.006	0.007	-0.019	1.000

Bayesian Analysis

Posterior Autocorrelations

Parameter	Lag 1	Lag 5	Lag 10	Lag 50
Intercept	0.7556	0.2216	0.0532	0.0187
Age	0.7657	0.2121	0.0626	-0.0074
BMI	0.7394	0.2074	0.0567	-0.0008
Dispersion	0.0348	0.0206	-0.0015	-0.0031

Geweke Diagnostics

Parameter	z	Pr > |z|
Intercept	-1.4414	0.1495
Age	0.4848	0.6278
BMI	1.1714	0.2414
Dispersion	0.4732	0.6361

Effective Sample Sizes

Parameter	ESS	Autocorrelation Time	Efficiency
Intercept	1424.4	7.0207	0.1424
Age	1305.5	7.6597	0.1306
BMI	1462.0	6.8400	0.1462
Dispersion	8252.7	1.2117	0.8253

Bayesian Analysis

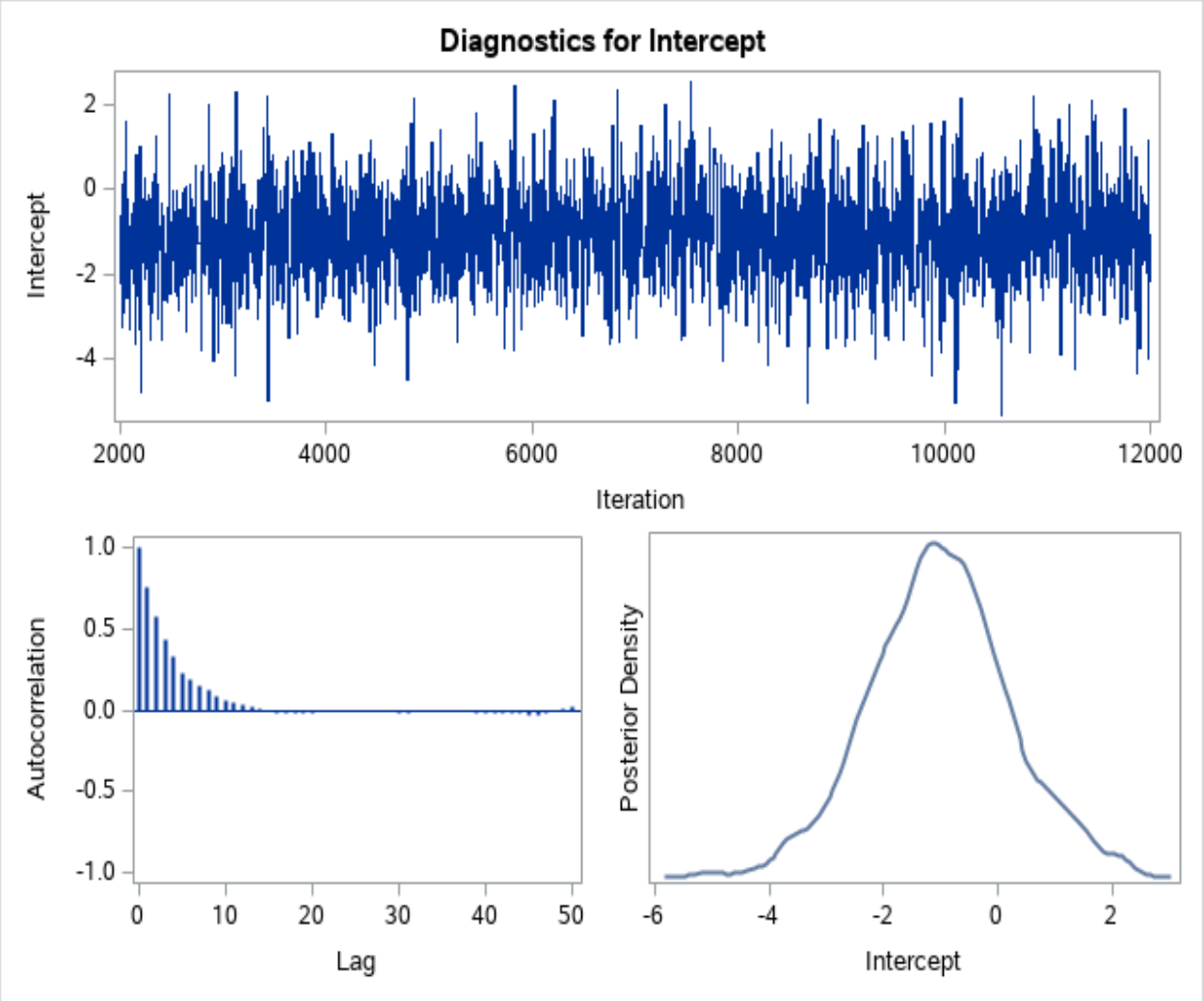

Figure 4: In the panel of the diagnostic plots above, the first graph shows good sharp spikes for the posterior distribution of the intercept. Autocorrelations are high in the first ten lags but low towards the end, and the posterior density is is approximately normal andnot too smooth.

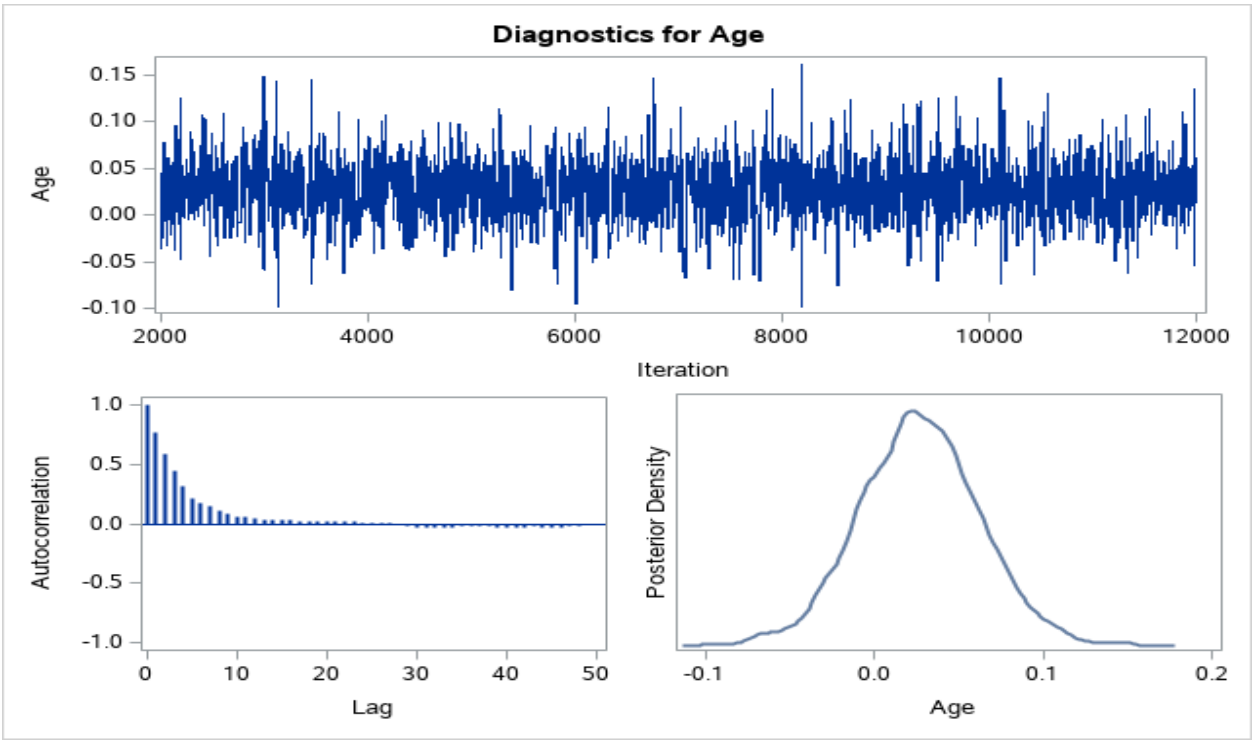

Figure 5: In the panel of the diagnostic plots above, the first graph shows sparsely good spikes for the posterior distribution of the Age. Autocorrelations are high in the first ten lags but low towards the end, and the posterior density is is approximately normal andnot smooth between -0.1 and 0.2.

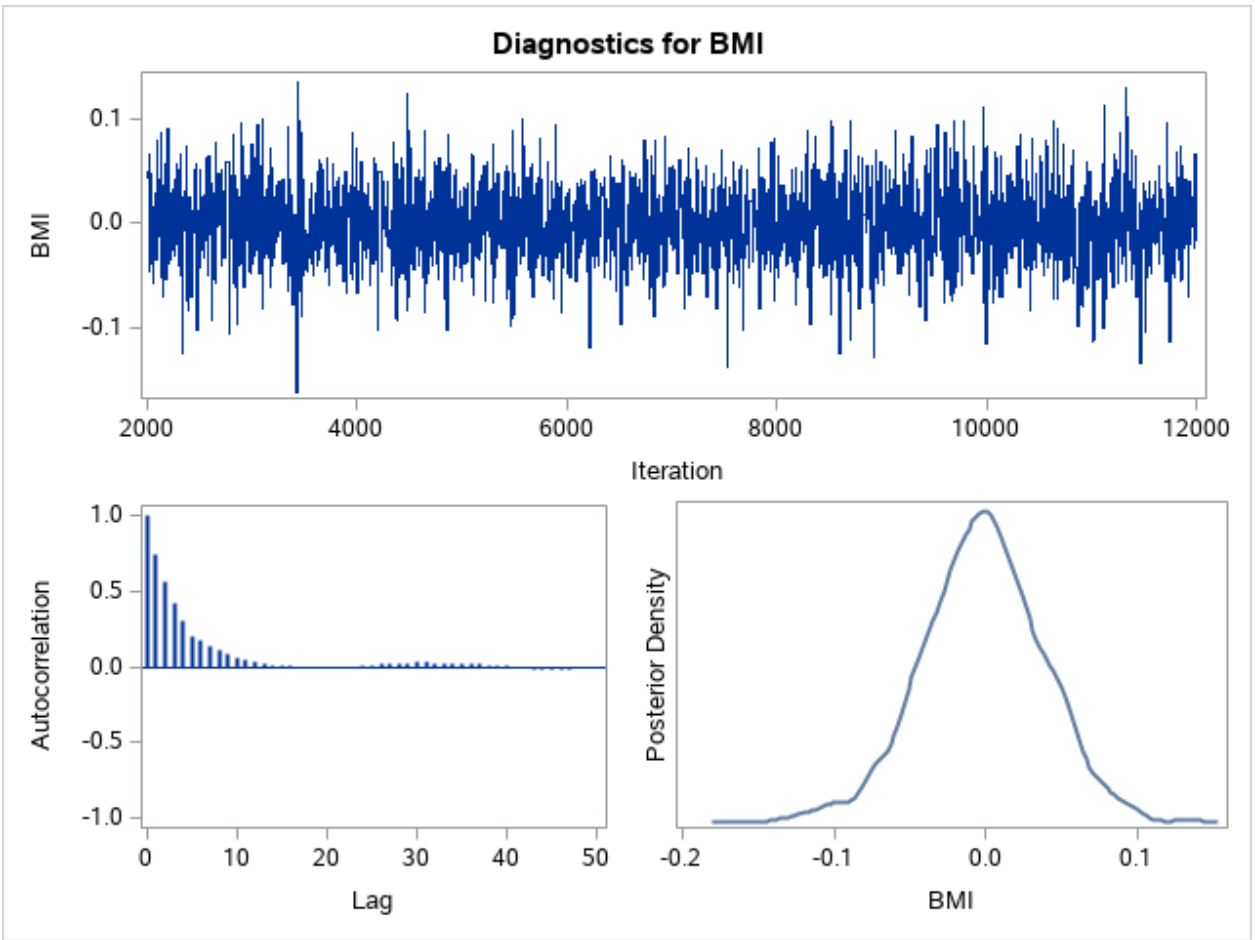

Figure 6: In the panel of the diagnostic plots above, the first graph shows sharpspikes for the posterior distribution of the Body Mass Index. Autocorrelations are high in the first ten lags but low towards the end, and the posterior density is is approximately normal and a bit smooth -0.1 and 0.1.

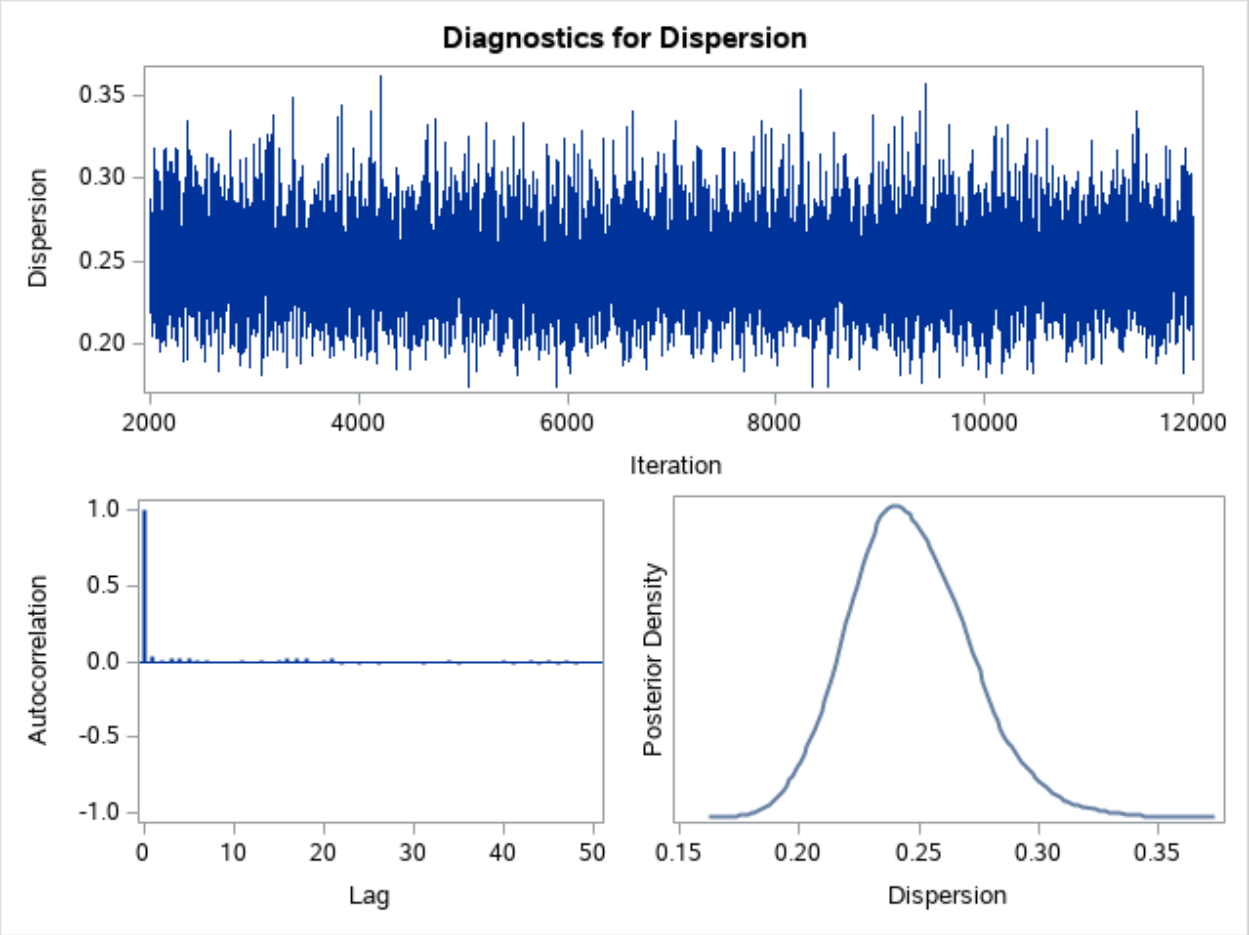

Figure 7: In the panel of the diagnostic plots above, the first graph shows densely good spikes for the posterior distribution of the dispersion. Autocorrelations are high in the first lag but low towards the end, and the posterior density is is approximately normal between 0.15 and 0.35.

Above are examples of important diagnostic plots automatically produced by GENMOD (and by other procedures in SAS) (Fig. 1, 2, 3, 4, 5, 6 and 7). These three plots are produced for each parameter in the model. Results for the intercept parameter is shown in Fig. 1 and 2. In each figure there are 3 sub-plots: trace plot, autocorrelation plot and posterior density plot. In Fig. 1, the first plot is the trace plot, a sequential graph of all the MCMC samples in the chain. Note that the first sample starts at 2000 because the burn-in was 2000 (default). There are then another 10,000 samples generated. The trace plot in Fig. 1 is the ideal that one should be looking forward for. There is no trend, and the values jump randomly above and below the central value. The autocorrelation plot shows the serial correlation of each sample in the chain with the previous samples; lag 1 is for each sample with the previous one, lag 2 is for each sample with the one two samples previously, and so on. At lag 0, the autocorrelation is 1, by definition, but is shown to provide perspective in interpretation. In Fig. 1, the lag 1 correlation is very close to 0, and the remaining correlations are even closer to 0. This is ideal. Finally, the posterior density plot is an estimate of the posterior distribution based on the 10,000 samples used from the trace plot (Fig. 1). A kernel smoothing algorithm is applied to the values to produce the curve. This posterior density is the ultimate goal of the Bayesian analysis. Summary statistics are derived from the samples that comprise the empirical density, such as the mean, median, standard deviation, and so on. Figure 7 shows an example that is slightly less than ideal. First, note that only 200 samples are generated in this example after the 2,000 *burn-in* samples and as such, the values are less crowded. One can see that the samples do not simply jump randomly around a central value, but sometimes move a bit slowly above and below the center value.

We know that positive values of beta are associated with increased probability that patients with chronic heart disease belong to the corresponding age and body mass index. Looking at the plots above, age and body mass index have positive beta values.

Negative values of beta are associated with decreased probability that patients with chronic heart disease belong to the corresponding age and body mass index. Looking at the plots above, none of the parameters fall in this group. The posterior distributions allow us to compute the mean and standard deviation of beta values per age and per BMI respectively as illustrated in the plots above. Based on these results, we can conclude that the beta posteriors are distributed uniformly across age and BMI respectively. This evidence is in support of the varying-intercept models.

Model	D-Optimality	A-Optimality	D-Efficiency	A-Efficiency
BayesLogit	2.8638e+011	1.341	99.806	99.683

IV. 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 using STATA 15. To evaluate the proposed criteria, an algorithm to evaluate the analytically intractable design criterion is derived. Bayesian logistic regression has the benefit that it gives us a posterior distribution rather than a single point estimate. The Bayes Logistic showed high D- and A-efficiency of 99.806 and 99.683 respectively and positive optimality. The results also showed that age and Body Mass Index positively affect the incidence of having chronic heart disease.

REFERENCES

[1] Ali, H. U., Lasisi K. E., 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.

Ali, H. is an academic staff in the department of Mathematics, Faculty of Natural Sciences, University of Jos. He holds Masters Degress in Statistics and currently undergoing his PhD in Biostatistics with bias Bayesian inference. He has published over five (5) articles in high impact factor journals, He is a member of many professional bobies amongst them are SIAM, PSSN, NMS etc.

Nwaosu, S. C. is a professor of Statistics in the department of Statistics, Joseph Sawuaan Tarka University of Agriculture, Makurdi, Nigeria. He has published over forty(40) articles in high impact factor journals, He is a member of many professional bodies amongst them are SIAM, PSSN, NMS etc.

Lasisi, K. E. is a professor of Statistics in the Department of Mathematical Sciences, Abubakar Tafawa Balewa University, Nigeria. He has published over forty(30) articles in high impact factor journals, He is a member of many professional bodies amongst them are SIAM, PSSN, NMS etc.

Abdulkadir, A. is a professor of Statistics in the Department of Mathematical Sciences, Abubakar Tafawa Balewa University, Nigeria. He has published over forty(30) articles in high impact factor journals, He is a member of many professional bodies amongst them are SIAM, PSSN, NMS etc.