Bayesian methods for clinicians

Razieh Bidhendi Yarandi[1,2], Kazem Mohammad[1], Hojjat Zeraati[1], Fahimeh Ramezani Tehrani[2],
Mohammad Ali Mansournia*[1]

Abstract

 Background: The Bayesian methods have received more attention in medical research. It is considered as a natural paradigm for dealing with applied problems in the sciences and also an alternative to the traditional frequentist approach. However, its concept is somewhat difficult to grasp by nonexperts. This study aimed to explain the foundational ideas of the Bayesian methods through an intuitive example in medical science and to illustrate some simple examples of Bayesian data analysis and the interpretation of results delivered by Bayesian analyses. In this study, data sparsity, as a problem which could be solved by this approach, was presented through an applied example. Moreover, a common sense description of Bayesian inference was offered and some illuminating examples were provided for medical investigators and nonexperts.

 Methods: Data augmentation prior, MCMC, and Bayes factor were introduced. Data from the Khuzestan study, a 2-phase cohort study, were applied for illustration. Also, the effect of vitamin D intervention on pregnancy outcomes was studied.

 Results: Unbiased estimate was obtained by the introduced methods.

 Conclusion: Bayesian and data augmentation as the advanced methods provide sufficient results and deal with most data problems such as sparsity.

Keywords: Bayesian inference, Prior information, MCMC, Data augmentation

Introduction

Bayesian methods, dating back to the 18th century, were introduced by the English statistician and philosopher Thomas Bayes in an essay released after his death (1, 2). Its application faced with some obstacles and controversies such as computational obstacles and prior selection challenge through decades (3-6). It starts with questing the degree of belief related to a concept and then updating it with the available evidence. The concept of Bayesian inference is intuitive as something you may be faced with in everyday life, which made it popular in all aspects of science (7). During the last few years, Bayesian application has increased drastically, offering pragmatic solutions for problems that a traditional approach fails to deal with (8). For instance, biased data may occur in some circumstances, including data sparsity, presence of outliers, censoring, and confounding variable selection. Bayesian methods have gained huge popularity due to the flexibility of modelling through a computer software to deal with these is-

_____

Corresponding author: Dr Mohammad Ali Mansournia, ansournia_m@sina.tums.ac.ir

1. Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
2. Reproductive Endocrinology Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

---

↑*What is "already known" in this topic:*
The increase in the use of Bayesian analysis in the medical researches has made it essential to present its complex concepts in a lucid language. To many nonexpert users, their computationally-intensive approaches have the form of a "black box".

→*What this article adds:*
This study aimed at offering a common sense description of Bayesian inference through an intuitive approach and providing some illuminating examples for medical investigators and nonexperts.

sues. Data sparsity occurs in the case of rare events or when some values of the predictor are associated with only 1 outcome value called separation (9); eg, when all cases of macrosomia occurred in women with gestational diabetes mellitus. In this case, estimated measures of association are subjected to bias with very big values and extraordinary wide ranges of 95% CI. The Bayesian approach can be considered as a remedial tool in this case (10-13).

Conceptually, now is not the era of starting everything from scratch; you may have "prior" knowledge about an investigation, as you know the probability of catching a cold in winter is higher. In this case, the question is to estimate a parameter, here probability of catching a cold, so you will first start collecting data to find the "likelihood" of catching a cold by objective evidence.

If you are a frequentist, you are done by only analyzing the obtained data ignoring any prior information. Contrary to frequentist approach, Bayesian method accounts the prior knowledge in the process of analysis to estimate "posterior" knowledge of the probability of catching cold.

Statistically speaking, by considering a prior distribution for the parameter of interest, the likelihood of data and Bayes rule, you can estimate the posterior distribution of the parameter.

Bayes Rule: Posterior distribution ∝ prior distribution× likelihood of data

In this study, it was aimed to provide a gentle technical introduction with an application to Bayesian inference useful for researchers in the medical fields. First, a short overview is presented regarding the history, concept, and definition of some jargons in this area of statistics and then its application is illustrated through an illuminating example along with its Stata software code to show how these ideas can be put into practice. Here, it is tried to make the concepts more tangible for nonexperts. Also, a neat catalogue of references in case readers are more interested in the bigger picture, more details, or a more advanced coverage of the topic is provided.

Methods
Bayesian approach in data analysis

Since its inception, there have been vigorous debates regarding the application of the Bayesian approach as an alternative to classical methods (4, 6, 14). The fundamental idea in Bayesian inference is to combine the prior knowledge and the available evidence. One can update one's knowledge when additional information becomes available. For example, you want to know the probability of breast cancer in a seemingly healthy woman (Fig. 1). Of course, the probability would be different if she were tested positive on a screening test. Bayes rule takes this extra information into account and shows how it should be evaluated. In this case, one would like to calculate posterior probability of having breast cancer given the positive test. Bayes rule combines the prior information of the positive test with the likelihood of a positive test given having the disease as below:

P(A): Probability of having the disease = 2%

P(B) = Probability of positive test = sum of probabilities of all independent ways to achieve a positive test = probability of true positive + probability of false positive = (having cancer × true positive) + (not having cancer × false positive) = (0.02 × .8) + (0.98 × 0.15) = 0.163

P (B | A) = probability of positive test given having the disease = 80%

P (A | B) = P (B | A) P(A)/P(B) = (.8×.02)/0.163 = 9.8%.

It shows that considering the prior information of the positive test increases the probability of diagnostic positive from 2% to 9.8%.

Calculating the posterior measures by classical Bayes rule was not that much straightforward, and in some cases needed intensive computation, which is why it was set aside for a long time until the advent of the computer (15). In this case, instead of calculating the posterior, methods were introduced to simulate it. The posterior simulation methods were evolved by the emergence of the MCMC approach in the 1990s, which is actually a computer-driven sampling approach. In this study, the mechanism of MCMC sampling was represented with Metropolis-Hastings algorithm through an example (Appendix). Generally speaking, prior information is also a key part of Bayesian methods and illustrates the knowledge about an uncertain parameter of estimate that is combined with the likelihood of data to provide the posterior measure (15, 16). Robust Bayesian analysis showed how much Bayesian answers are sensitive to uncertain inputs. Under certain conditions, which is large sample size, the prior is dominated by the likelihood of data, so evaluation of the true prior may not be required. In this regard, the prior effect is consistent with a large set of informative and diffuse priors. On the other hand, a Bayesian method that utilizes
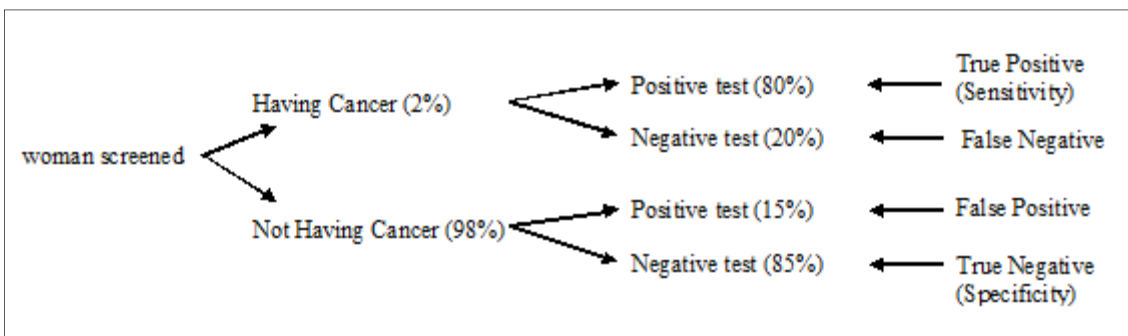

Fig. 1. Diagram represents sensitivity, specificity, true positive, and true negative

weakly informative priors to quantify the sensitivity of parameters to sparse data was proposed. Parameter estimates will be shrunk to the prior mean in case of the sparseness. In practice, to check the influence of prior distributions, the sensitivity analysis can be applied, which compares the posterior inferences under various reasonable selections of prior distributions (17-20). The interpretation of posterior results can also be very straightforward. Posterior credible interval and Bayes factor are introduced instead of 95% confidence interval and p value. Also, 95% credible interval shows there is a probability of 0.95 that Bayesian calculated parameter of interest is within this set, which is precisely like what is sometimes wrongly interpreted as frequentist estimates. While true frequentist interpretation of 95% CI is the interval which will contain the true value on 95% of occasions if a study were repeated many times using samples from the same population.

Data augmentation prior approach

An easy and more tangible method of data augmentation prior (DAP) was introduced as an alternative to Bayesian methods (10-12). Its concept is truly straightforward. First, translate prior to a so-called pseudo data, and then add this amount of information to the real dataset; then, ordinary methods of calculating parameters of interest can be applied. Contrary to the Bayesian approach whose concept is like a black-box and in some cases takes much time to gain results, DAP provides a more understandable and applicable estimation process. We presented the result of an illustrative example by DAP method as well, and the results are virtually identical (21).

Below proved an illustration of DAP method by a simple example:

Consider that in a clinical trial LnOR (Variance of lnOR) for an event were estimated as LnOR = 2.0 (Variance of lnOR =1.0).

The prior information for OR with 95% limits between $\frac{1}{4}$ and 4 was obtained from a meta-analysis. Mean and variance of prior for Ln (OR) are estimated as follow:

$prior\ mean\ Ln(OR) = average\ of\ 95\%\ limits$

$$= \frac{(Ln\left(\frac{1}{4}\right) + Ln(4))}{2} = 0$$

$prior\ variance\ Ln(OR)$
$$= (\frac{Width\ of\ interval\ in\ Ln(RR)units}{Width\ of\ interval\ in\ standard\ devision\ units})^2$$
$$= (\frac{\left|Ln\left(\frac{1}{4}\right) - Ln(4)\right|}{2 * 1.96})^2 = 0.5$$

Therefore, a normal prior with 0 mean and 0.5 variance was defined. The contribution of prior and data information to estimate posterior mean and variance could be assessed through their inverse variances equaling $\frac{1}{Variance\ of\ prior} = \frac{1}{0.5} = 2$ and $\frac{1}{variance\ of\ LnOR} = \frac{1}{1} = 1$, showing that prior information dominated the data information by nearly 2 times. Posterior mean and variance for Ln (OR) could be estimated as the following weighted

averaging rule of thumb; posterior mean for

$$ln(OR) = \frac{\frac{Mean\ of\ Prior}{Variance\ of\ prior} + \frac{lnOR}{Variance\ of\ LnOR}}{\frac{1}{Variance\ of\ prior} + \frac{1}{Variance\ of\ LnOR}} = \frac{\frac{0}{0.5} + \frac{2}{1}}{\frac{1}{0.5} + \frac{1}{1}} = 0.67,$$

posterior variance for

$$ln(OR) \approx \frac{1}{\frac{1}{Variance\ of\ prior} + \frac{1}{Variance\ of\ LnOR}} = \frac{1}{\frac{1}{0.5} + \frac{1}{1}} = 0.33,$$

and 95% posterior CI for
$OR \approx$

$$exp\left(posterior\ mean \pm 1.96 * (Posterior\ variance)^{\frac{1}{2}}\right) = exp\left(0.67 \pm 1.96 * (0.33)^{\frac{1}{2}}\right) = exp(-0.47, 1.80) = (0.63, 11.94).$$

The width of CI obtained from this method was narrower than the ordinary approach. The "Penlogit" is a Stata command for DAP (22).

Bayes factor approach

Bayes factor is a Bayesian approach to hypothesis testing. It measures the strength of evidence by providing a number for quantifying the evidence in favor of a scientific hypothesis. To clarify more, Bayes factor compares the relative likelihood given to both null and alternative hypotheses, while p value is estimated based on just the null hypothesis (H0) (23). By Bayes rule, posterior odds in favor of H0 would be the prior odds multiplied by likelihood ratio, namely, the Bayes factor, where Bayes factor equaled the ratio of likelihood of data under the null hypothesis to the likelihood of data under the alternative hypothesis. In fact, prior odds are transformed to posterior odds via Bayes factor, which presents how a belief is changed by data. In other words, Bayes factor is the evidence for the alternative hypothesis versus the null hypothesis. Bayes factor equal 1 provides no evidence, 1/3 − 1: anecdotal evidence for H1, 1/3 − 1/10: moderate evidence for H1, 1/10 − 1/30: strong evidence for H1, 1/30 − 1/100: very strong evidence for H1, < 1/100: extreme evidence for H1 (23-26).

Illustrative example

Effect of vitamin D intervention on neonatal death: Data were extracted from the Khuzestan study, a 2-phase cohort study. The effect of vitamin D intervention on pregnancy outcomes was studied, and the details of the study protocol have previously been published (27). The contingency table shows a cell has only one observation, which is subjected to data sparsity. In this case, the ordinary logistic regression fails to truly calculate the odds ratio and 95% CI. The big odds ratio of neonatal death and the unusually wide confidence interval represent the bias estimation of the measures OR = 125 (95% CI: 15.8 to 1000). The data were reanalyzed by the Bayesian and DAP approaches with various priors, ranging from strong (low variance) to weak prior information for logistic regression coefficient.

A common distribution for regression coefficient is normal distribution (28). The mean of 0 and variance of

Results

Table 1 shows the median of prior (95% prior limits). Results revealed a reasonable shrink of ORs and 95% credible intervals, especially for stronger prior (Table 2), which is in line with the literature (29, 30). The results for Bayesian and DAP approaches are almost the same. In addition, to test the null hypothesis (H_0: Beta = 0), Bayes factor approach was applied, and the result showed that the evidence did not worth more than a bare mention in favor of H_0. Therefore, vitamin D intervention significantly decreased the odds of neonatal death. Stata codes for the Bayesian and DAP approaches are available in the Appendix.

Discussion

The volume of literature published on Bayesian inference has proved its popularity among medical studies (31-41). In this study, it was aimed to present some Bayesian approaches in an intuitive language for clinicians. Bayesian and DAP methods were run on data of Khuzestan co-

0.5 were considered to 10 provide various degree of beliefs regarding the magnitude of the relationship between exposure and disease.

hort study, which were subjected to sparse data problem to illustrate the application. DAP and MCMC both provided almost the same result, while generally DAP is a more understandable and convenient method, especially for nonexperts (42). Generally, Bayesian and data augmentation as the advanced methods provide sufficient results, unbiased estimates, and deal with most data problems such as sparsity when traditional frequentist method fails.

Conclusion

This study demonstrated the foundational concepts of Bayesian inference through an intuitive illustration for nonexperts.

Conflict of Interests
The authors declare that they have no competing interests.

Table 1. Different classes of Prior information with plausible ranges: median (95%limit)

Priors	Exact prior median OR	95% prior limit OR
Normal (0, 0.5)	1	(1/4, 4)
Normal (0, 1)	1	(1/7, 7)
Normal (0, 1.38)	1	(1/10, 10)
Normal (0, 2)	1	(1/16, 16)
Normal (0, 10)	1	(1/492, 492)

Table 2. Intervention and Neonatal Death contingency table and results of logistic regression

		No-Intervention	Intervention
Neonatal death	Yes	15	1
	No	97	787

Ordinary Logistic Regression Approach

	Beta	SE	OR=exp (Beta)	95% Confidence Interval (P-value)
	4.82	.032206	125	15.8 to1000 (P=0.000)

MCMC Approach

Prior	Posterior Mean of Beta	$^\epsilon$MCSE	OR=exp (Beta)	95% Credible Interval (P-value)
Normal (0, 0.5)	2.5	.011877	12.2	5.7 to 27.1
Normal (0, 1)	3.2	.016299	23.8	9.1 to 62.5
Normal (0, 1.38)	3.4	.018704	29.9	9.8 to 85.4
Normal (0, 2)	3.6	.016885	36.5	11.1 to 122.6
Normal (0, 10)	4.7	.031016	112.4	22.2 to 909

DAP Approach

Prior	Penalized Beta	SE	OR=exp (Beta)	95% Credible Interval (P-value)
Normal (0, 0.5)	2.5	.4226359	12.5	5.5 to 28.8
Normal (0, 1)	3.1	.5032273	22.7	8.3 to 62.5
Normal (0, 1.38)	3.4	.5473388	28.9	9.8 to 84.4
Normal (0, 2)	3.6	.602612	37.5	11.5 to 121.6
Normal (0, 10)	4.8	.8479396	118.8	20.1 to 909

Bayes Factor Approach

Prior odds for null hypothesis	*Bayes factor	Posterior odds for null hypothesis	Evidence
1	0.0000031	0.0000031	Evidence not worth more than a bare mention in favor of H_0

$^\epsilon$ Monte Carlo estimation of standard error for Beta regression coefficient
* The ratio of likelihood of data in model with intercept effect versus model with intercept and vitamin D-intervention effects

References

1. Fienberg SE, When did Bayesian inference become" Bayesian"? Bayesian Anal. 2006;1(1):1-40.
2. Bayes T. Essay towards solving a problem in the doctrine of chances. Biometrika. 1958;45:293-315.
3. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis. 2013: Chapman and Hall/CRC.
4. Efron B. Why isn't everyone a Bayesian? Am Stat. 1986;40(1):1-5.
5. Berger J. The case for objective Bayesian analysis. Bayesian Anal. 2006;1(3):385-402.
6. Bland JM, Altman DG. Bayesians and frequentists. BMJ. 1998;317(7166):1151-1160.
7. Armitage P, Berry G, Matthews JNS. Stat Methods Med Res. 4th Edition ed. 2001: Wiley-Blackwell.
8. Gatsonis C, Kass RE, Carlin B, Carriquiry A, Gelman A, Verdinelli I, et al, (Eds.) Case studies in bayesian statistics. Springer Science & Business Media. 2012:83.
9. Mansournia MA, Geroldinger A, Greenland S, Heinze G. Separation in logistic regression: causes, consequences, and control. Am J Epidemiol. 2018;187(4):864-70.
10. Greenland S, Christensen R. Data augmentation priors for Bayesian and semi-Bayes analyses of conditional-logistic and proportional-hazards regression. Stat Med. 2001;20(16):2421-2428.
11. Greenland S, Mansournia MA. Penalization, bias reduction, and default priors in logistic and related categorical and survival regressions. Stat Med. 2015;34(23):3133-3143.
12. Greenland S, Mansournia MA, Altman DG. Sparse data bias: a problem hiding in plain sight. BMJ. 2016;352.
13. van Zwet E. A default prior for regression coefficients. Stat Methods Med Res. 2019;28(12):3799-807.
14. Berger J. The case for objective Bayesian analysis. Bayesian Anal. 2006;1(3):385-402.
15. Armitage P, Colton T. Encyclopedia of Biostatistics. Wiley. 2005;8.
16. Gelman A. Prior distribution. Encyclopedia of environmetrics, 2006. 4.
17. Berger JO, Moreno E, Pericchi LR, Bayarri MJ, Bernardo JM, Cano JA, et al. An overview of robust Bayesian analysis. Test. 1994;3(1):5-124.
18. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis. CRC press; 2013 Nov 1.
19. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, et al. Global sensitivity analysis: the primer. John Wiley & Sons; 2008 Feb 28.
20. Hamra GB, MacLehose RF, Cole SR. Sensitivity analyses for sparse-data problems-using weakly informative bayesian priors. Epidemiology (Cambridge, Mass.), 2013;24(2): 233-239.
21. Greenland S. Bayesian perspectives for epidemiological research: I. Foundations and basic methods. Int J Epidemiol. 2006;35(3):765-775.
22. Discacciati A, Orsini N, Greenland S. Approximate Bayesian logistic regression via penalized likelihood by data augmentation. Stat J. 2015;15(3):712-736.
23. Goodman SN. Introduction to Bayesian methods I: measuring the strength of evidence. Clin Trials. 2005;2(4):282-90.
24. Goodman SN. Toward evidence-based medical statistics. 2: The Bayes factor. Ann Intern Med. 1999;130(12):1005-13.
25. Agresti A. Categorical Data Analysis. 3rd ed. 2013: Wiley.
26. Lavine M, Schervish MJ. Bayes factors: What they are and what they are not. Am Stat. 1999;53(2):119-122.
27. Rostami M, Ramezani Tehrani F, Simbar M, Bidhendi Yarandi R, Minooee S, Hollis BW, et al. Effectiveness of prenatal vitamin D deficiency screening and treatment program: a stratified randomized field trial. J Clin Endocrinol Metab. 2018;103(8):2936-2948.
28. Gelman A, Jakulin A, Pittau MG, Su YS. A weakly informative default prior distribution for logistic and other regression models. Ann Appl Stat. 2008;2(4):1360-83.
29. Bhimji KM, Naburi H, Aboud S, Manji K. Vitamin D Status and Associated Factors in Neonates in a Resource Constrained Setting. Int J Pediatr. 2018;2018.
30. McGuigan I, Flatley C, Kumar S. Maternal vitamin D levels and the risk of perinatal death. J Matern Fetal Neonatal Med. 2017;30(13):1544-1548.
31. Mohammadian-Khoshnoud M, Moghimbeigi A, Faradmal J, Yavangi M. A two-component Bayesian mixture model to identify implausible gestational age. Med J Islam Repub Iran. 2016;30:440.
32. Pourhoseingholi A, Baghestani AR, Ghasemi E, Akbarzadeh Baghban A, Ghazarian M. Bayesian Zero- Inflated Poisson model for prognosis of demographic factors associated with using crystal meth in Tehran population. Med J Islam Repub Iran. 2018;32(1):136-141.
33. Ebrahim Valojerdi A, Janani L. A brief guide to propensity score analysis. Med J Islam Repub Iran. 2018;32(1):717-720.
34. Gelman A, Rubin DB. Markov chain Monte Carlo methods in biostatistics. Stat Methods Med Res. 1996;5(4):339-355.
35. Hammersley JM, Morton KW. Poor Man's Monte Carlo. J R Stat Soc Ser A Stat Soc. Series B. 1954;16(1): 23-38.
36. Hamra G, MacLehose R, Richardson D. Markov Chain Monte Carlo: an introduction for epidemiologists. Int J Epidemiol. 2013;42(2):627-634.
37. Smith AFM, Gelfand AE. Bayesian Statistics without Tears: A Sampling-Resampling Perspective. Am Stat. 1992;46(2):84-88.
38. Turner BM, Van Zandt T. A tutorial on approximate Bayesian computation. J Math Psychol. 2012;56(2): 69-85.
39. Witmer J. Bayes and MCMC for Undergraduates. Am Stat. 2017;71(3):259-264.
40. Kruschke J. Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan. 2014: Academic Press.
41. Friston K. The history of the future of the Bayesian brain. Neuroimage Clin. 2012;62(2):1230-1233.
42. Sullivan SG, Greenland S. Bayesian regression in SAS software. Int J Epidemiol. 2012;42(1):308-317.
43. Gilks WR, Spiegelhalter D, Richardson S. Markov Chain Monte Carlo in Practice. 1996: Chapman & Hall/CRC.
44. Hesterberg T. Weighted average importance sampling and defensive mixture distributions. Technometrics. 1995;37:185-194.
45. Zhang J, Fu S. An efficient Bayesian uncertainty quantification approach with application to k-ω-γ transition modeling. Comput Fluids. 2018;161:211-224.
46. Gilks WR, Wild P. Adaptive Rejection Sampling for Gibbs Sampling. J R Stat Soc Ser C Appl Stat. 1992;41(2):337-348.
47. Robert C, Casella G. Introducing Monte Carlo Methods with R. 2010: Springer-Verlag New York.
48. Brooks SG, Jones A, Meng GL, Meng XL. Handbook of Markov Chain Monte Carlo. 2011 Chapman & Hall/CRC.

Appendix

MCMC history and algorithm: MCMC consists of 2 parts: The Monte Carlo method and the Markov chain stochastic process (3, 34, 43). Historically, the Monte Carlo method is the first approach to simulate measures such as mean or variance. It uniformly generates samples from a probability function and numerically calculates the measures. This sampling approach was improved by some methods such as importance sampling (44, 45) and rejection sampling (46). As a special case of importance sampling, weighted-prior, instead of uniformly generating samples from the prior, assigns a weight obtained from the likelihood of data to each sample. Therefore, the posterior is obtained through weighted-prior samples. MCMC, as a dynamic and advanced approach, is more practical (36). There are 2 popular algorithms for MCMC method: Metropolis-Hastings first introduced by Metropolis in 1953, and its special case Gibbs sampler introduced in 1984 (47). Recent developments have provided an extensive literature. Armitage has provided a neat catalogue of the references and summaries (7, 15, 48). To present its algorithm, imagine a researcher is interested in estimating the mean of birth weight in infants with GDM mothers to test whether they are subjected to fetal macrosomia. The researcher knows that the weight is normally distributed with a standard deviation of 100. She has only observed a weight of an infant equal to 3500 grams and wants to apply MCMC approach to draw samples from the "target" distribution, Bayesian talking "posterior", which represents the probability of each possible value of the population mean given this single observation, normal ($\mu = 3500$, $\sigma = 100$). To draw samples from the distribution of weight, MCMC firstly starts with an educated guess. Suppose this initial guess is 3600; then, a chain of new samples is created by this initial sample. Two steps are considered in the process of generating a new sample: first, adding a small random noise to the initial and generating a "proposal" for the new sample; second, deciding whether it is an appropriate sample or not through an acceptance rule. Moreover, there are various ways of creating proposals and rules for accepting or rejecting this candidate. Metropolis-Hastings is one of the famous methods. The following illustration is the Metropolis-Hasting algorithm to generate a chain of samples:

1. Generate an initial sample for weight, eg, 3600.
2. Generate a proposal sample by adding a random noise to the initial sample from normal ($\mu = 0$, $\sigma = 50$) proposal distribution, eg, 3650.
3. It is time to decide whether to accept the proposal as the next sample or not. Compare the height of the posterior at the value of the new proposal against the height of the posterior at the initial sample. If this proportion is greater than one, one is considered as the value.
4. Generate a pseudo-random number from a uniform (0, 1) distribution called u.
5. If this proportion is greater than u, accept the new proposal; otherwise, reject it with a probability equal to the value of the proportion of the heights. In the case of rejecting the proposal, the initial is selected as the new sample again and the process of sampling is iterated until enough samples are generated.

In Bayesian jargon, considering the proposal as prior distribution and target as prior multiplied by likelihood which is posterior distribution.

After generating samples from the posterior distribution, computational methods can be applied to estimate the mean, median, and mode of infants' weight. The credible interval is also estimable for the mean of posterior which, unlike 95% CI, showed the range of the estimated parameter.

Stata codes for Bayesian and DAP methods of analysis for estimating the effect of history of GDM on fetal macrosomia

Bayesian:
```
bayesmh neonatalDeath interv, likelihood(logit) prior({interv}, normal(0,1)) prior({_cons}, flat)
bayesmh neonatalDeath interv, likelihood(logit) prior({interv}, normal(0,10)) prior({_cons}, flat)
```

DAP:
```
#penlogit neonatalDeath interv, nprior(interv ln(1) 1) or
#penlogit neonatalDeath interv, nprior(interv ln(1) 10) or
```