Comparative Analysis of two 1 × J Contingency Tables using Bayesian Perspective

Debodeep Banerjee[#], Trina Sahoo[#], M. Subbiah[*], M. R. Srinivasan[$]

[#]Post Graduate, Department of Statistics, University of Madras

[*]Assistant professor, Department of Mathematics, Presidency College, Chennai-5)

[$]Professor and Head, Department of Statistics, University of Madras

Abstract.

The wide spread use of Bayesian Inference is mainly attributed to the scope of researchers' ability to incorporate prior distribution for the parameters, surge in statistical computations and model developments in multiple disciplines. This study has exploited these aspects in to categorical data of two one-dimensional variables each with same J levels of categories. The notion focussing on ratio of independent Beta distribution is considered using closed form approach with Gauss hypergeometric function and Monte-Carlo techniques. Entire approach is illustrated with a primary data set that aims to study the impact of gender on perceived important social issues; essential computations are carried out using R program. The modelling advantages of Bayesian approach has been studied and the results are directly interpreted regarding the context of the problem.

***Keywords**:* Bayesian Inference, Multinomial distribution, Contingency table, Ratio of Beta variables, Monte Carlo integration, R.

Contents

Chapter 1

Introduction 2

Chapter 2

Bayesian Inference and an overview 3

Chapter 3

Multinomial distribution 6

Chapter 4

Application and analysis 13

Chapter 5

Discussion and conclusion 29

References 31

1. Introduction

Bayesian inference is a method of statistical inference in which Bayes' theorem is used to update the knowledge as more information becomes available. Bayesian method provides a complete paradigm for both statistical inference and decision making under uncertainty. Anthony O'Hagan has emphasized the principles of Bayesian statistics as prior knowledge, subjective probability and so on. However, Bayesian inference is very much relevant in the case of categorical data analysis also. In fact, it enables us to collect more information from a categorical data by the virtue of its operational leniency and efficiency. This work is focused on the comparative study of two 1×J Contingency table with cell counts said to follow a multinomial distribution. The notion focusing on the ratio of two independent Beta distribution is considered using closed form of Gauss hypergeometric function. Pham-Jia (2000) emphasized the discussion on ratio between a pair of independent beta ratio. The prior distribution is defined for the data set as Dirichlet distribution, and with a reasonable likelihood function, inference about parameters of interest is obtained from marginal posterior distributions as Beta distribution and thereafter summarizing the posterior distribution.

This article deals with the worrying issues of our country. We selected few issues like Terrorism, Unemployment, Corruption, Environmental Threats, Female Security, Moral Decline, Religious Extremism and Price Hike. Moreover, we split our observations with respect to their genders, i.e. two sets of samples were collected. One represented the male population and the other come from the female. Individuals can select exactly one issue which they consider to be the most worrying issue and we collect the data via social media. Our data model is multinomial dealing with 8 variables. In Bayesian inference our main objective is to estimate the unknown parameter (θ). Here the parameters signify the proportion of individual supporting an issue. We first set our multinomial distribution $(x \mid \theta)$ and based on the range of the unknown parameter we set the distribution for the parameter (θ). Thereby, we set our prior distribution. By Bayesian paradigm, Posterior is prior times likelihood. Thus, if the probability $\theta \mid x$ is the posterior distribution. For a given data we often write:

$$P(\theta \mid x) \propto P(x \mid \theta) f(\theta)$$

From the posterior distribution, we obtain the Bayesian Point estimate, i.e., mean, variance and credible interval and we obtain the Jeffys prior. Then we calculate the marginal distribution of the posterior distribution to determinate the hyper-parameter. We are interested in comparative study of the parameters to examine which issue is more sensible to be considered as most worrying issue. In order to fulfil the interest, we go for ratio test.

There are little hardships in the numerical evaluation of the integrations. In that case, we took help of R-language to get the proper result. The necessary R codes have been provided in the appendix portion.

2. Bayesian Inference and an overview

2.1 Introduction

By Inference, we mean the reasoning involved in drawing a conclusion or making a logical judgment on the basis of circumstantial evidence. Statistics in a wider perspective can be classified as frequentist and Bayesian paradigms with conceptual and computational difference between the two methodologies. The fundamental difference between frequentist and Bayesian is related with nature of the unknown parameter in the model. In the Bayesian view the unknown parameters are treated as random variable with known distributions; the conditional distribution of unknowns given known follow from applying Bayes theorem to the model specifying the joint distribution of known and unknown quantities (Rubin,1984). In the frequentist view parameters are treated as fined random quantity.

Bayesian probability is an interpretation of concept of probability, in which, instead of frequency or propensity of some phenomenon, probability is interpreted as reasonable expectation representing a state of knowledge or a quality of personal belief (Berger, 2006). The major drawbacks which could be cited are computational difficulties and the formal use of subjective probability in application of Bayesian inference over the many years. However, uses of statistical software and computational approaches such as Markov chain Monte Carlo (MCMC) method have highly influenced Bayesian approach (Raftery and Lewis, 1996). However, this article is limited with Bayesian methods and the three fold objectives are understanding the theoretical basis, methodology and computation involved in the Bayesian learning; the entire exercise is illustrated using multinomial model dealing with categorical data.

2.2 Bayesian inference

A statistical model consists of the observations of a random variable X distributed according to $f(x \mid \theta)$ with parameter θ unknown and belongs to the parameter space Θ of finite dimension. Uncertainty on the parameter could be modeled through a probability distribution p(θ) called a prior distribution. The inference is then based on θ conditioned on X.

The mechanism of Bayesian approach involves three basic steps:

1) Assign distribution for all unknown parameters based on the range of the parameters.
2) Define the likelihood function of the data given the parameter, i.e., $L(x \mid \theta)$.
3) Determine the posterior distribution of the parameter, i.e., $\pi(\theta \mid x)$.

Literature in constructing prior distributions provide following broad approaches:-
The knowledge of prior can be extended to understand the various term found in literature in general:
The prior may be determined subjectively which represent beliefs (Berger). The experimenter may have information or data that can be used to help formulate a prior known as objective and informative prior. This could take atleast two forms:

a) Historical data on the distribution of parameter values.
b) Data from experiments done prior to the one being undertaken.

A non-informative prior is one which is dominated by the likelihood function. In other words, such a prior does not change much over the region in which the likelihood is appreciable and does not assume large values outside that region (Box and Tiao). When prior distributions have no population basis, it is difficult to construct such a prior distribution which plays a minimal role in the posterior distribution. Such distributions are sometimes called reference prior distribution (Gelman,1995).
A class of prior distributions is said to form a conjugate family if the posterior density π(θ/x) is in that class for all x whenever the prior density is in the class. In other words, when a family of *conjugate* prior exists, choosing a prior from that family simplifies calculation of posterior distribution, i.e., the posterior and prior distribution are similar. There is another objective or non-informative prior probability

distribution known as Jeffreys prior which is proportional to the square root of the determinant of the Fisher information matrix, i.e., $p(\theta) \propto \sqrt{|J(\theta)|}$. Parameters of prior distributions are a kind of *hyperparameter*.

The knowledge of parameter θ can be updated using Bayes theorem as

$\pi(\theta/x) = \frac{f(\theta,x)}{f(x)} = \frac{p(\theta) f(x/\theta)}{f(x)}$ where $f(x) = \int p(\theta).f(x/\theta)d\theta$ in case of continuous data and $f(x) = \sum p(\theta)\, f(x/\theta)$ in case of discrete data and the resulting density $\pi(\theta/x)$ is called posterior distribution. Hence f(x) is independent of θ. We can conclude that posterior is prior times likelihood, i.e.,

$\pi(\theta \mid x) \propto f(x \mid \theta) \times p(\theta)$

The posterior probability distribution contains all the current information about the parameter θ. The desirable numerical summaries of the distribution include mean, median, mode, credible interval. For example, the summary of posterior distribution are helpful in practice especially a point estimate in θ is it's posterior mean as

$$\hat{\theta} = E(\theta/x) = \int \theta\, \pi(\theta/x)d\theta$$

If the posterior distribution is skewed, then the posterior mode can be used. The variance of the posterior density is called posterior variance or posterior standard deviation that describes the uncertainty in the parameter can be used to characterize the dispersion of parameter θ.

$$\sigma_\theta = \sqrt{var(\theta/x)} = \sqrt{\int (\theta - \hat{\theta})^2 \pi(\theta/x)d\theta}$$

A credible interval of the posterior probability corresponds to the domain of the posterior probability distribution or predictive distribution used for interval estimation. The $(1-\alpha)\%$ is credible interval is defined as $(q_{\alpha/2}, q_{1-\alpha/2})$ where the $q_{\alpha/2}$ represents $\frac{\alpha}{2}$quantile of the posterior distribution.

Let x|θ be the data density and p(θ) be the prior for θ and the posterior for θ be π(θ|x),then we can predict the future probability which is known as **Predictive Probability**.

For example, consider a Binomial distribution with number of success x=10 and number of trail n=14.
The probability mass function is given by,

$$f(x \mid \theta) = \binom{14}{10} \theta^{10} (1-\theta)^4 ; x = 0,1,2,\ldots,0 < \theta < 1$$

By the prior knowledge before observing the data, we set Beta distribution of 1st kind with two positive shape parameter a and b as our prior distribution.
Then the probability density function is given by,

$$p(\theta) = \frac{1}{B(a,b)} \theta^{a-1} (1-\theta)^{b-1} ; 0 < \theta < 1, a > 0, b > 0$$

Here a and b are hyperparameter.
Now,

$$f(x \mid \theta) p(\theta) = \binom{14}{10} \frac{1}{B(a,b)} \theta^{10+a-1} (1-\theta)^{4+b-1}$$

The posterior distribution will be

$$\pi(\theta \mid x) = \frac{\theta^{10+a-1} (1-\theta)^{4+b-1}}{B(10+a, 4+b)}$$

Therefore, $\theta \mid x \sim Beta(10+a, 4+b)$

This is an example of Conjugate prior distribution. The posterior thus obtained is a closed form and hence summaries are direct.

Summarization:

$$Mean = \frac{10+a}{14+a+b}$$

$$Variance = \frac{(10+a)(4+b)}{(14+a+b)^2(15+a+b)}$$

The 95% posterior interval for which P($\theta<\theta_l$)=P($\theta>\theta_u$)=0.025. Together provide, $p(\theta_l < \theta < \theta_u) = 1-\alpha$

Precisely, if Y ~ Beta (a,b) then

$$F = \frac{b}{a}\frac{Y}{1-Y} \sim \text{F(2a, 2b)}$$

This is useful to compute probabilities such as $p(\theta < \theta_l) = p\left(\frac{b}{a}\frac{\theta}{1-\theta} < \frac{b}{a}\frac{\theta_l}{1-\theta_l}\right)$ and subsequently probabilities can be obtained by referring F-table.

Table 1: Comparison of posterior summary for the data n=14 and x=10 under three priors

Prior	Posterior Mean	Posterior Variance	Credible Interval
Jeffreys Beta(0.5,0.5)	0.7	0.013125	(0.45,0.895)
Haldane(approx.) beta(0.01,0.01)	0.71	0.013596	(0.46,0.91)
Bayes-Laplace beta(1,1)	0.6875	0.012637	(0.45,0.88)

If we fail to get a closed form we can go for computation with R and MC.

3. Multinomial distribution

3.1 The multinomial distribution

In the previous chapter, we have discussed on the notion of a binomial distribution. The binomial distribution can be generalised to a **Multinomial distribution** which allows more than two possible outcomes. For example, if we consider the role of a die, there are 6 possible outcomes. So, the event of rolling of a die (once or more than once) can be probabilistically explained and analysed using a Multinomial distribution.

A random vector $\underset{\sim}{X}$= (X_1, X_2 , ……., X_j) is said to follow a Multinomial distribution with parameters N and $\theta_1, \theta_2, \ldots, \theta_j$ if its probability mass function is given by,

$$\mathbf{f(X|\theta)} = \frac{N_1!}{X_1\,!\,X_2!\,X_3!\ \ldots\ X_j!} \times (\boldsymbol{\theta_1}^{x_1}\ \ \boldsymbol{\theta_2}^{x_2}\ \ \boldsymbol{\theta_3}^{x_3}\ \ \ldots\ \boldsymbol{\theta_j}^{x_j}\ )$$

$$\textbf{Where, } \sum_{i=1}^{j} \boldsymbol{x_i} = \boldsymbol{N_1},\ \ \sum_{i=1}^{j} \boldsymbol{\theta_i} = \mathbf{1} \textbf{ and } \mathbf{0} \le \boldsymbol{\theta_i} \le \mathbf{1},\ \mathbf{i = 1,2,3,\ldots,j}$$

Here, $\underset{\sim}{X}$ denotes the vectors of counts of the number of observation and θ_i being the probability of the occurrence of the i^{th} event. The sum of the respective probabilities is restricted to 1. i.e. $\sum_{i=1}^{j} \theta_i = 1$.

3.2 Construction of prior

To construct the prior distribution, we consider the parameter(s) of the prior distribution as the random variable(s) of the prior distribution. Likewise, in the present scenario, θ_i's will be treated as random variables in the prior distribution. With range, 0 to 1.

In chapter 2, we have seen that the prior distribution of a binomial distribution is a beta distribution. In the multinomial set up, the prior distribution is multivariate generalization of the beta distribution which is also known as Dirichlet distribution. In both cases (Binomial and Multinomial) the priors are actually conjugate priors.

The prior distribution will be,

$(\theta_1, \theta_2, \theta_3, \ldots, \theta_j) \sim dirichlet(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_j)$

The probability density function will be,

$$p(\theta) = \frac{\Gamma(\alpha)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)\ldots\ldots\Gamma(\alpha_j)} \times (\theta_1^{\alpha_1 - 1}\theta_2^{\alpha_2 - 1}\theta_3^{\alpha_3 - 1} \ldots\ \theta_j^{\alpha_j - 1}) \quad , \alpha_i \ge 0,\ \ 0 \le \theta_i \le 1$$

3.3 Construction of posterior

Hence, the posterior will be,

$\Pi(\theta|x)\ \alpha\ p(\theta) \times f(x|\theta)$

Hence,

$\theta|X \sim diri(\alpha_1 + x_1, \alpha_2 + x_2, \ldots, \alpha_j + x_j)$

<u>*Determination of the values of* α_i</u>

$(X_1, X_2, X_{3,}, \dots, X_j) \sim$ Multinomial $(N_1, \theta_{1,} \theta_{2,} \theta_{3,} \dots, \theta_j)$

Then,

$$L(X|\theta) = \frac{N_1!}{X_1!\, X_2!\, X_3! \,\ldots\ldots X_j!} \times (\theta_1{}^{x_1}\ \theta_2{}^{x_2}\ \theta_3{}^{x_3}\ \ldots \theta_j{}^{x_j}\ )$$

$$= (N_1! / \textstyle\prod_{i=1}^{j} x_i\,!) \times \prod_{i=1}^{j-1} \theta_i{}^{x_i} \times\ (1 - \sum_{i=1}^{j-1} \theta_i)^{N_1 - \sum_{i=1}^{j-1} x_i}$$

Where, $\sum_{i=1}^{j} x_i = N_1,\ \ \sum_{i=1}^{j} \theta_i = 1$

$=> \log(L(\theta|X) = constant + \sum_{i=1}^{j} x_i\ \ ln\, \theta_i + (N_1 - \sum_{i=1}^{j-1} x_i)\ln(1 - \sum_{i=1}^{j-1} \theta_i)$

So,

$$\frac{\partial l}{\partial \theta_1} = \frac{x_1}{\theta_1} - \frac{N_1 - \sum_{i=1}^{j-1} x_i}{1 - \sum_{i=1}^{j-1} \theta_i}$$

$$\frac{\partial^2 l}{\partial \theta_1^2} = -\frac{x_1}{\theta_1^2} - \frac{N_1 - \sum_{i=1}^{j-1} x_i}{1 - \sum_{i=1}^{j-1} \theta_i}$$

Following the above traits, we can claim for i=1,2,3, …, j-1 that:

$$\frac{\partial l}{\partial \theta_i} = \frac{x_i}{\theta_i} - \frac{N_1 - \sum_{i=1}^{j-1} x_i}{1 - \sum_{i=1}^{j-1} \theta_i}$$

$$\frac{\partial^2 l}{\partial \theta_i^2} = -\frac{x_i}{\theta_i^2} - \frac{N_1 - \sum_{i=1}^{j-1} x_i}{1 - \sum_{i=1}^{j-1} \theta_i}$$

Again, we have already got:

$$\frac{\partial l}{\partial \theta_i} = \frac{x_i}{\theta_i} - \frac{N_1 - \sum_{i=1}^{j-1} x_i}{1 - \sum_{i=1}^{j-1} \theta_i}$$

$$\frac{\partial^2 l}{\partial \theta_i \partial \theta_k} = -\frac{N_1 - \sum_{i=1}^{j-1} x_i}{(1 - \sum_{i=1}^{j-1} \theta_i)^2}$$

$where\ i, k = 1,2, \dots, j-1,\ i \neq k$

Now,

$$E\left(\frac{\partial^2 l}{\partial \theta_i^2}\right) = \frac{E(X_i)}{\theta_i^2} - \frac{N_1 - \sum_{i=1}^{j-1} E(x_i)}{(1 - \sum_{i=1}^{j-1} \theta_i)^2}$$

$$= \frac{N_1 \theta_i}{\theta_i^2} - \frac{N_1 - N_1 \sum_{i=1}^{j-1} \theta_i}{(1 - \sum_{i=1}^{j-1} \theta_i)^2}$$

$$= \frac{N_1}{\theta_i} - \frac{N_1}{1 - \sum_{i=1}^{j-1} \theta_i}$$

Again,

$$E\left(\frac{\partial^2 l}{\partial\theta_i\partial\theta_k}\right) = -\frac{N_1 - \sum_{i=1}^{j-1} E(x_i)}{1 \times \left(1 - \sum_{i=1}^{j-1}\theta_i\right)^2} = -\frac{N_1 - N_1\sum_{i=1}^{j-1}\theta_i}{1 \times \left(1 - \sum_{i=1}^{j-1}\theta_i\right)^2} = -\frac{N_1}{1 - \sum_{i=1}^{j-1}\theta_i}$$

Hence,

$$|J| = \begin{vmatrix} \frac{N_1}{\lambda_1} + \frac{N_1}{1-\sum_{i=1}^{j-1}\lambda_i} & \frac{N_1}{1-\sum_{i=1}^{j-1}\lambda_i} & \cdots & \frac{N_1}{1-\sum_{i=1}^{j-1}\lambda_i} \\ \frac{N_1}{1-\sum_{i=1}^{j-1}\lambda_i} & \frac{N_1}{\lambda_2} + \frac{N_1}{1-\sum_{i=1}^{j-1}\lambda_i} & \cdots & \frac{N_1}{1-\sum_{i=1}^{j-1}\lambda_i} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{N_1}{1-\sum_{i=1}^{j-1}\lambda_i} & \frac{N_1}{1-\sum_{i=1}^{j-1}\lambda_i} & \cdots & \frac{N_1}{\lambda_{j-1}} + \frac{N_1}{1-\sum_{i=1}^{j-1}\lambda_i} \end{vmatrix}$$

$$= \frac{N^{j-1}}{\lambda_1 \ldots \lambda_{j-1}(1-\sum_{i=1}^{j-1}\lambda_i)}$$

Now, $p(\theta)\ \alpha\ |J(\theta)|^{\frac{1}{2}}$

$$= \prod_{i=1}^{j-1}\theta_i^{-1/2}\ \left(1 - \sum_{i=1}^{j-1}\theta_i\right)^{-1/2} = \prod_{i=1}^{j-1}\theta_i^{\frac{1}{2}-1}\ \left(1 - \sum_{i=1}^{j-1}\theta_i\right)^{\frac{1}{2}-1}$$

Therefore, $\alpha_1 = \alpha_2 = \alpha_3 = \ldots = \alpha_j = ½$

If $(X_1, X_2, X_3, \ldots\ldots, X_j)$ ~ Multinomial $(N_1, \theta_1, \theta_2, \theta_3, \ldots\ldots, \theta_j)$

Then, θ~ diri (1/2, 1/2, 1/2,….., 1/2)

And, θ|X ~ diri (α_1+x_1, α_2+x_2, α_3+x_3, ………., α_j+x_j)

Jeffry's prior for θ is,

$\alpha_i = 1/2$, i=1,2,3,…,j

$$(\theta_1, \theta_2, .., \theta_j | X) \sim Diri(x_1+\alpha_1, x_2+\alpha_2, x_3+\alpha_3, x_4+\alpha_4, x_5+\alpha_5, ..., x_j+\alpha_j)$$

3.4 Marginal distribution of theta

$$f(\theta|X) = \frac{\Gamma(N_1+\alpha)}{\prod_{i=1}^{j}\Gamma(x_i+\alpha_i)}\prod_{i=1}^{j}\theta_i^{x_i+\alpha_i-1}, \quad \sum_{i=1}^{j}(x_i+\alpha_i) = N_1+\alpha$$

$$\iiint\underset{\sum\theta_i\leq 1}{\iiint}\int \frac{\Gamma(N_1+\alpha)}{\prod_{i=1}^{j}\Gamma(x_i+\alpha_i)}\prod_{i=1}^{j}\theta_i^{x_i+\alpha_i-1}\ \partial\theta_1.....\partial\theta_j$$

$$= \iiint\underset{\sum\theta_i\leq 1}{\iiint} \frac{\Gamma(N_1+\alpha)}{\prod_{i=1}^{j}\Gamma(x_i+\alpha_i)}\prod_{i=1}^{j}\theta_i^{x_i+\alpha_i-1}(1-\theta_1-\theta_2-.....-\theta_{j-1})^{x_8+\alpha_8-1}\ \partial\theta_2.....\partial\theta_{j-1}$$

Consider, $\theta_{j-1} = (1-\theta_1-\theta_2-\theta_3-\ldots-\theta_{j-2})\, u_1$

$$\partial\theta_{j-1} = (1-\theta_1-\theta_2-\theta_3-\ldots-\theta_{j-2})\,\partial u_1$$

Then the integral will be,

$$\iiint\underset{\sum\theta_i\le 1}{\iiint} \frac{\Gamma(N_1+\alpha)}{\prod_{i=1}^{j}\Gamma(x_i+\alpha_i)}\theta_1^{x_1+\alpha_1-1}\theta_2^{x_2+\alpha_2-1}\theta_3^{x_3+\alpha_3-1}\ldots\theta_{j-2}^{x_{j-2}+\alpha_{j-2}-1}(1-\theta_1-\theta_2-\theta_3-\ldots-\theta_{j-2})^{x_{j-1}+\alpha_{j-1}}u_1^{x_{j-1}+\alpha_{j-1}-1}$$

$$\times[(1-\theta_1-\theta_2-\theta_3-\ldots-\theta_{j-2})-(1-\theta_1-\theta_2-\theta_3-\ldots-\theta_{j-2})u_1]^{x_j+\alpha_j-1}\partial u_1\partial\theta_2\partial\theta_3\ldots\ \partial\theta_{j-2}$$

$$=\int\underset{\sum\theta_i<1}{\iint}\iint\frac{\Gamma(N_1+\alpha)}{\prod_{i=1}^{j}\Gamma(x_i+\alpha_i)}\theta_1^{x_1+\alpha_1-1}\theta_2^{x_2+\alpha_2-1}\theta_3^{x_3+\alpha_3-1}\ldots.\theta_{j-2}^{x_{j-2}+\alpha_{j-2}-1}(1-\theta_1-\theta_2-\theta_3-\ldots-\theta_{j-2})^{x_{j-1}+\alpha_{j-1}+x_j+\alpha_j-1}$$

$$\times\,[\int_0^1 u_1^{x_{j-1}+\alpha_{j-1}-1}(1-u_1)^{x_j+\alpha_j-1}\partial u_1\partial\theta_2\partial\theta_3\ldots\partial\theta_{j-2}$$

$$=\int\underset{\sum\theta_i<1}{\iint}\iint\frac{\Gamma(N_1+\alpha)}{\prod_{i=1}^{j}\Gamma(x_i+\alpha_i)}\theta_1^{x_1+\alpha_1-1}\theta_2^{x_2+\alpha_2-1}\theta_3^{x_3+\alpha_3-1}\ldots\theta_{j-2}^{x_{j-2}+\alpha_{j-2}-1}(1-\theta_1-\theta_2-\theta_3-\ldots-\theta_{j-2})^{x_{j-1}+\alpha_{j-1}+x_j+\alpha_j-1}$$

$$\times\ B(x_{j-1}+\alpha_{j-1},x_j+\alpha_j)\ \partial\theta_2\partial\theta_3\ldots\partial\theta_{j-2}$$

Proceeding in this way, we get,

$$\prod_1(\theta_1\,|\,X)=\frac{\Gamma(N_1+\alpha)}{\Gamma(x_1+\alpha_1)\Gamma(x_2+\alpha_2+x_3+\alpha_3+\ldots+x_j+\alpha_j)}\quad\times\quad\theta_1^{x_1+\alpha_1-1}(1-\theta_1)^{\sum_{j=2}^{j}x_j+\alpha_j-1}\qquad, j=2,\ldots,j$$

$$\therefore\ (\theta_1|X)\sim Beta(\,x_1+\alpha_1, N_1+\alpha-x_1-\alpha_1)$$

We can claim,

$$\left(\theta_j\middle|X\right)\sim Beta\left(\,x_j+\alpha_j, N_1+\alpha-x_j-\alpha_j\right)$$

Posterior Mean and Variance.

$$E(\theta_j|X) = \frac{\alpha_j+x_j}{N_1+\alpha}$$

$$V(\theta_j|X) = \frac{(x_j+\alpha_j)(N_1+\alpha-x_j-\alpha_j)}{(N_1+\alpha)^2(N_1+\alpha+1)}$$

3.5 Ratio of two independent beta distribution

Pham-Jia (2000) introduced the concept of ratio of two independent beta distribution of first kind. Suppose X and Y be two independently distributed random variables with beta distribution of 1st kind. The challenge was to determine the distribution of X/Y. We try to apply this concept under the light of Bayesian paradigm.

Consider the standard beta distribution with density,

X~Beta(a,b)

Y~Beta(α,β)

X and Y are independent.

$$f_1(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} \text{ , } 0<x<1$$

$$= 0 \text{ , otherwise}$$

$$f_2(y) = \frac{1}{B(\alpha,\beta)} y^{\alpha-1}(1-y)^{\beta-1}, \; 0<y<1$$

$$= 0, \text{ otherwise}$$

Since X and Y are independent, the joint pdf of X and Y will be,

$$f(x,y) = \frac{1}{A} x^{a-1}(1-x)^{b-1} y^{\alpha-1}(1-y)^{\beta-1} \text{ , } 0<x<1 \text{ ; } 0<y<1$$

Where, A=B(a, b).B(α, β). Let ,

$$W = \frac{x}{y}$$

$V = y$ so that

$$0 \le y < 1$$
$$\therefore 0 \le v < 1$$
$$0 \le x < 1$$
$$0 \le wv < 1$$

But $w \ge 0$

$0 < v < 1$ and $w > 0$

Then $x = wv$ and $y = v$

Jacobian of the transformation is:

$$|J| = \frac{\partial(x,y)}{\partial(w,v)} = \begin{vmatrix} \frac{\partial x}{\partial w} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial w} & \frac{\partial y}{\partial v} \end{vmatrix}$$

$$= \begin{vmatrix} v & w \\ 0 & 1 \end{vmatrix}$$

$$= v$$

The joint PDF of w and v is

$$\Phi(w,v)=\frac{1}{A}(wv)^{a-1}v^{\alpha-1}(1-wv)^{b-1}(1-v)^{\beta-1}v$$

$$=\frac{1}{A}w^{a-1}v^{a-1}v^{\alpha-1}(1-wv)^{b-1}(1-v)^{\beta-1}v$$
$$=\frac{1}{A}w^{a-1}v^{a+\alpha-2+1}(1-wv)^{b-1}(1-v)^{\beta-1}$$
$$=\frac{1}{A}w^{a-1}v^{a+\alpha-1}(1-wv)^{b-1}(1-v)^{\beta-1}$$

$w>0$, $0<v<1$ and $wv\le 1$

Or equivalently the joint PDF of w and y is:

$$\Phi(w,y)=\frac{1}{A}w^{a-1}y^{a+\alpha-1}(1-wy)^{b-1}(1-y)^{\beta-1}$$

If w<1,

$$f(w)=\int_0^1 \phi(w,y)dy$$

$$=\frac{w^{a-1}}{A}\int_0^1 y^{a+y-1}(1-wy)^{b-1}(1-y)^{\beta-1}dy \qquad \text{(i)}$$

Consider the Gauss hypergeometric function in three parameters.

$${}_2F_1(m;n;k;x)=\frac{1}{B(m,k-m)}\int_0^1 u^{m-1}(1-u)^{k-m-1}(1-xu)^{-n}du \qquad \text{(ii)}$$

Comparing (i) and (ii),

$$a+\alpha=m$$
$$k-m=\beta$$
$$\Rightarrow k=m+\beta=a+\alpha+\beta$$
$$n=1-b$$

$\therefore$ equation (i) becomes,

$$f(w)=\frac{w^{a-1}.B(a+\alpha,\beta)}{A}{}_2F_1(a+\alpha;1-b;\beta+\alpha+a;w)$$

If $w\ge 1$, $let\ wy=t$

$$y=\frac{t}{w}$$
$$dy=\frac{dt}{w}$$
$$f(w)=\frac{w^{a-1}}{A}\int_0^1\left(\frac{t}{w}\right)^{a+\alpha-1}(1-t)^{b-1}\left(1-\frac{t}{w}\right)^{\beta-1}\frac{dt}{w}$$

$$=\frac{w^{a-1}}{A}\left(\frac{1}{w}\right)^{a+\alpha}\int_0^1 t^{a+\alpha-1}(1-t)^{b-1}\left(1-\frac{t}{w}\right)^{\beta-1}dt$$

$$=\frac{w^{-(\alpha+1)}}{A}\int_0^1 t^{a+\alpha-1}(1-t)^{b-1}\left(1-\frac{t}{w}\right)^{-(1-\beta)}dt$$

$$=\frac{w^{-(\alpha+1)}.B(a+\alpha,\beta)}{A}{}_2F_1(a+\alpha;1-\beta;a+b+\alpha;\frac{1}{w})$$

Hence the pdf of $w=\frac{x}{y}$ is,

$$f(w)=\begin{cases}\frac{w^{a-1}.B(a+\alpha,\beta)}{A}{}_2F_1(a+\alpha;1-b;a+\alpha+\beta;w), 0<w<1\\ \frac{w^{-(1+\alpha)}.B(a+\alpha,b)}{A}{}_2F_1(a+\alpha;1-\beta;a+\alpha+b;w), w\geq 1\end{cases}$$

4. Application and analysis

4.1 Discussion on data

This chapter explicitly deals with an application of what we have discussed so far in the last three. Chapters. Here, we name the problem as **“Most worrying issue in India”.** A questionnaire was formed with 8 specific options which was set focussed on India’s socio-political and economic scenario. The respondents were requested to opt exactly one of the given options. It was an online based survey and two target population was fixed, viz. males and females. The samples were independent to each other. We classified the samples with respect to their genders and thus formed two different response tables. One is solely based on the male responses and another one was based on the female responses. Hence, we get two specific multinomial models. The objective of this survey was to check which issue is more concerning to which genders. i.e. gender wise consciousness towards a specific issue was the main objective of the experiment.

We start with analysing the male responses and necessary summary measures have been calculated including their credibility intervals. Similar exercise has been done with the female responses also. Later, the notion of **ratio of two independent beta variables** is applied to draw relevant inferences.

Table 2: A snapshot of the responses for the questionnaire related to the data analysis and ‘.’ Indicates the choice of the responses.

Sample	Most Worrying Issue in India							
	Terrorism	Unemployment	Corruption	Environmental Threats	Moral Decline	Religion Extremism	Female Security	Price Hike
M01						•		
M02						•		
M03						•		
M04							•	
M05			•					
M06					•			
M07			•					
M08						•		
M09		•						
M10						•		
M11					•			
M12						•		
M13	•							
M14						•		
M15					•			
M16					•			
M17				•				
M18						•		
M19	•							
M20						•		
.								
.								
.								
M59			.					
Total	2	10	12	5	7	18	3	2
			Grand Total= 59					

4.2.2 Data distribution

$(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)$ ~Multinomial $(N, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8)$

The probability mass function will be,

$$f(X|\theta) = \frac{N!}{X_1!\,X_2!\,X_3!\,X_4!\,X_5!\,X_6!\,X_7!\,X_8!} \times ({\theta_1}^{x_1}\ {\theta_2}^{x_2}\ {\theta_3}^{x_3}\ {\theta_4}^{x_4}\ {\theta_5}^{x_5}\ {\theta_6}^{x_6}\ {\theta_7}^{x_7}\ {\theta_8}^{x_8})$$

Where, $\sum_{i=1}^{8} x_i = N$, $\sum_{i=1}^{8} \theta_i = 1$ and $0 \le \theta_i \le 1$, i = 1,2,3,…,8

θ_i's are the parameters of the distribution.

If we put the values of $X_1, \ldots, X_8$, the pmf becomes,

$$f(X|\theta) = \frac{59!}{2!\ 10!\ 12!\ 5!\ 7!\ 18!\ 3!\ 2!} \times ({\theta_1}^2{\theta_2}^{10}{\theta_3}^{12}{\theta_4}^5{\theta_5}^7{\theta_6}^{18}{\theta_7}^3{\theta_8}^2),\ \sum_{i=1}^{8} \theta_i = 1 \text{ and } 0 \le \theta_i \le 1,\ i=1,2,\ldots,8$$

4.2.3 Construction of prior

To construct the prior distribution, we consider the parameter(s) of the prior distribution as the random variable(s) of the prior distribution. Likewise, in the present scenario, θ_i's will be treated as random variables in the prior distribution. With range, obviously, 0 to 1.

The prior distribution will be,

$(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8) \sim dirichilet(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8)$

The probability density function will be,

$$p(\theta) = \frac{\Gamma(\alpha)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)\Gamma(\alpha_4)\Gamma(\alpha_5)\Gamma(\alpha_6)\Gamma(\alpha_7)\Gamma(\alpha_8)} \times \quad \theta_1^{\alpha_1 - 1} \ldots\ldots \theta_8^{\alpha_8 - 1},\ \alpha_i \ge 0,\ 0 \le \theta_i \le 1$$

4.2.4 Determining posterior

Hence, the posterior will be,

$\Pi(\theta|x)\ \alpha\ p(\theta) \times f(x|\theta)$

Hence,

$\theta|X \sim$ diri$(\alpha_1+x_1, \alpha_2+x_2, \alpha_3+x_3, \alpha_4+x_4, \alpha_5+x_5, \alpha_6+x_6, \alpha_7+x_7, \alpha_8+x_8)$

<u>Determination of the values of α_i</u>

$(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)$ ~ Multinomial $(N, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8)$

Then,

$$L(X|\theta) = \frac{N!}{X_1!\,X_2!\,X_3!\,X_4!\,X_5!\,X_6!\,X_7!\,X_8!} \times ({\theta_1}^{x_1}\ {\theta_2}^{x_2}\ {\theta_3}^{x_3}\ {\theta_4}^{x_4}\ {\theta_5}^{x_5}\ {\theta_6}^{x_6}\ {\theta_7}^{x_7}\ {\theta_8}^{x_8})$$

$$= (N!/\prod_{i=1}^{8} x_i!) \times \prod_{i=1}^{7} {\theta_i}^{x_i} \times (1 - \sum_{i=1}^{7} \theta_i)^{N - \sum_{i=1}^{7} x_i}$$

Where, $\sum_{i=1}^{8} x_i = N$, $\sum_{i=1}^{8} \theta_i = 1$

=> log$(L(\theta|X) = constant + \sum_{i=1}^{7} x_i \ \ln \theta_i + (N - \sum_{i=1}^{7} x_i)ln(1 - \sum_{i=1}^{7} \theta_i)$ = l (say)

So,

$$\frac{\partial l}{\partial \theta_1} = \frac{x_1}{\theta_1} - \frac{N - \sum_{i=1}^{7} x_i}{1 - \sum_{i=1}^{7} \theta_i}$$

$$\frac{\partial^2 l}{\partial \theta_1^2} = -\frac{x_1}{\theta_1^2} - \frac{N - \sum_{i=1}^{7} x_i}{(1 - \sum_{i=1}^{7} \theta_i)^2}$$

Following the above traits, we can claim for i=1,2,3,…, 7 that:

$$\frac{\partial l}{\partial \theta_i} = \frac{x_i}{\theta_i} - \frac{N - \sum_{i=1}^{7} x_i}{1 - \sum_{i=1}^{7} \theta_i}$$

$$\frac{\partial^2 l}{\partial \theta_i^2} = -\frac{x_i}{\theta_i^2} - \frac{N - \sum_{i=1}^{7} x_i}{(1 - \sum_{i=1}^{7} \theta_i)^2}$$

Again, we have already got:

$$\frac{\partial l}{\partial \theta_i} = \frac{x_i}{\theta_i} - \frac{N - \sum_{i=1}^{7} x_i}{1 - \sum_{i=1}^{7} \theta_i}$$

$$\frac{\partial^2 l}{\partial \theta_i \partial \theta_j} = -\frac{N - \sum_{i=1}^{7} x_i}{(1 - \sum_{i=1}^{7} \theta_i)^2}$$

$where\ i, j = 1,2,3, \ldots, 7,\ i \neq j$

Now,

$$E\left(\frac{\partial^2 l}{\partial \theta_i^2}\right) = \frac{E(X_i)}{\theta_i^2} - \frac{N - \sum_{i=1}^{7} E(x_i)}{(1 - \sum_{i=1}^{7} \theta_i)^2}$$

$$= \frac{N\theta_i}{\theta_i^2} - \frac{N - N\sum_{i=1}^{7} \theta_i}{(1 - \sum_{i=1}^{7} \theta_i)^2}$$

$$= \frac{N}{\theta_i} - \frac{N}{1 - \sum_{i=1}^{7} \theta_i}$$

Again,

$$E\left(\frac{\partial^2 l}{\partial \theta_i \partial \theta_j}\right) = -\frac{N - \sum_{i=1}^{7} E(x_i)}{1 * (1 - \sum_{i=1}^{7} \theta_i)^2} = -\frac{N - N\sum_{i=1}^{7} \theta_i}{1 * (1 - \sum_{i=1}^{7} \theta_i)^2} = -\frac{N}{1 - \sum_{i=1}^{7} \theta_i}$$

Hence,

$$|J| = \begin{vmatrix} \frac{N}{\theta_1} + \frac{N}{1-\sum_{i=1}^{7}\theta_i} & \frac{N}{1-\sum_{i=1}^{7}\theta_i} & \cdots & \frac{N}{1-\sum_{i=1}^{7}\theta_i} \\ \frac{N}{1-\sum_{i=1}^{7}\theta_i} & \frac{N}{\theta_2} + \frac{N}{1-\sum_{i=1}^{7}\theta_i} & \cdots & \frac{N}{1-\sum_{i=1}^{7}\theta_i} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{N}{1-\sum_{i=1}^{7}\theta_i} & \frac{N}{1-\sum_{i=1}^{7}\theta_i} & \cdots & \frac{N}{\theta_7} + \frac{N}{1-\sum_{i=1}^{7}\theta_i} \end{vmatrix}$$

$$= \frac{N^7}{\theta_1 \ldots \theta_7 (1-\sum_{i=1}^{7}\theta_i)}$$

Now, $p(\theta)\ \alpha\ \ |J(\theta)|^{\frac{1}{2}}$

$$= \prod_{i=1}^{7} \theta_i^{-1/2}\ \ (1-\sum_{i=1}^{7}\theta_i)^{-1/2} = \prod_{i=1}^{7} \theta_i^{\frac{1}{2}-1}\ \ (1-\sum_{i=1}^{7}\theta_i)^{\frac{1}{2}-1}$$

Therefore, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = ½$

If $(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)$ ~ Multinomial $(59, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8)$

Then, θ~ diri (1/2, 1/2, 1/2, 1/2, 1/2, 1/2, 1/2, 1/2)

And, θ|X ~ diri(2.5, 10.5, 12.5, 5.5, 7.5, 18.5, 3.5, 2.5)

Jeffrys prior for θ is,

$\alpha_i = 1/2$, i=1,2,3,…,8

4.2.5 Summary

- *Calculation of the marginals from a Dirichlet distribution.* *

$$(\ \ \theta_1, \theta_2, .., \theta_8 | X) \sim Diri(x_1+\alpha_1, x_2+\alpha_2, x_3+\alpha_3, x_4+\alpha_4, x_5+\alpha_5, x_6+\alpha_6, x_7+\alpha_7, x_8+\alpha_8)$$

$$\Pi(\theta|X) = \frac{\Gamma(N+\alpha)}{\prod_{i=1}^{8}\Gamma(x_i+\alpha_i)} \prod_{i=1}^{8} \theta_i^{x_i+\alpha_i-1}, \quad \sum_{i=1}^{8}(x_i+\alpha_i) = N+\alpha$$

Now, the marginal distribution of $(\theta_1|X)$ will be,

$$(\theta_1|X) \sim Beta(\ x_1+\alpha_1, N+\alpha-x_1-\alpha_1)$$

• Calculation of posterior mean and variance.

Table 3: The following table exhibits the posterior mean variance and the credible intervals of the marginal distributions of θ_i's based on the male responses

Random Variable	Expectation	Variance
$(\theta_1\|X) \sim$ Beta(2.5,60.5)	$E(\theta_1\|X)=0.039$	$V(\theta_1\|X)=0.000595$
$(\theta_2\|X) \sim$ Beta(10.5,52.5)	$E(\theta_2\|X)=0.166$	$V(\theta_2\|X)=0.00217$
$(\theta_3\|X) \sim$ Beta(12.5,50.5)	$E(\theta_3\|X)=0.198$	$V(\theta_3\|X)=0.00248$
$(\theta_4\|X) \sim$ Beta(5.5,57.5)	$E(\theta_4\|X)=0.087$	$V(\theta_4\|X)=0.00124$
$(\theta_5\|X) \sim$ Beta(7.5,55.5)	$E(\theta_5\|X)=0.119$	$V(\theta_5\|X)=0.001638$
$(\theta_6\|X) \sim$ Beta(18.5,44.5)	$E(\theta_6\|X)=0.293$	$V(\theta_6\|X)=0.00324$
$(\theta_7\|X) \sim$ Beta(3.5,59.5)	$E(\theta_7\|X)=0.055$	$V(\theta_7\|X)=0.00081$
$(\theta_8\|X) \sim$ Beta(2.5,60.5)	$E(\theta_8\|X)=0.039$	$V(\theta_8\|X)=0.00595$

* *Refer to page 9*

4.2.6 Credible intervals

Credibility intervals of the parameters. [1]

$$\theta \sim B(\alpha,\beta), then,$$
$$\frac{\beta}{\alpha}\frac{\theta}{1-\theta} \sim F_{m,n}$$
$$where, m = 2\alpha; n = 2\beta$$

Let the level of significance is 95%

Credible interval for $\theta_1|X$.

$$\theta_1|X \sim Beta(x_1+\alpha_1, N+\alpha-x_1-\alpha_1)$$
$$\theta_1|X \sim Beta(2.5, 60.5)$$

$$\frac{60.5}{2.5}\frac{\theta_1}{1-\theta_1} \sim F_{5,121}$$

$$Now, P(\theta_l \le \theta_1 < \theta_u) = 0.95$$

$$\therefore P(\theta_1 \le \theta_l) = 0.025$$

$$and, P(\theta_1 > \theta_u) = 0.025$$

$$Now,, P(\theta_1 > \theta_u) = 0.025$$

$$we\ can\ say, P\left(\frac{60.5}{2.5}\frac{\theta_1}{1-\theta_1} > \frac{60.5}{2.5}\frac{\theta_u}{1-\theta_u}\right) = 0.025$$

$$\Rightarrow P(F > F_u) = 0.025$$

$$\Rightarrow \frac{60.5}{2.5}\frac{\theta_u}{1-\theta_u} = 2.673$$

$$\Rightarrow \frac{\theta_u}{1-\theta_u} = 0.110$$

$$\Rightarrow \ \theta_u = 0.099$$

$$Again, P(\theta_1 \leq \theta_l) = 0.025$$

$$\Rightarrow P\left(\frac{60.5}{2.5}\frac{\theta_1}{1-\theta_1} \leq \frac{60.5}{2.5}\frac{\theta_l}{1-\theta_l}\right) = 0.025$$

$$\Rightarrow P(F < F_l) = 0.025$$

$$\Rightarrow \frac{60.5}{2.5}\frac{\theta_l}{1-\theta_l} = 0.164$$

$$\Rightarrow \frac{\theta_l}{1-\theta_l} = 0.0067$$

$$\Rightarrow \theta_l = 0.0066$$

Hence the credible interval for $\theta_1|X$ *is 0.099 and 0.0066.*

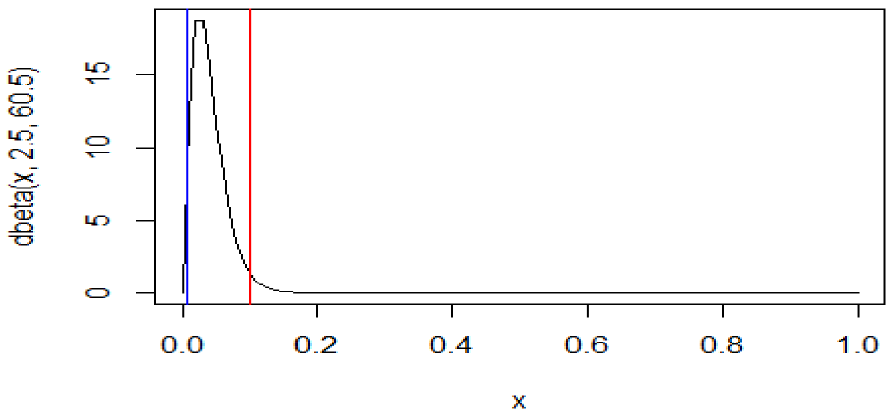

Diagram 01: Graphical representation for Credible Interval of $\theta_1|X$

Credible interval for $\theta_2|X$

$$\theta_2|X \sim Beta(x_2 + \alpha_2, N + \alpha - x_2 - \alpha_2)$$

$$\theta_2|X \sim Beta(10.5, 52.5)$$

$$\frac{52.5}{10.5}\frac{\theta_2}{1-\theta_2} \sim F_{21,105}$$

$$Now, P(\theta_l \leq \theta_2 < \theta_u) = 0.95$$

$$\therefore P(\theta_2 \leq \theta_l) = 0.025$$

$$and, P(\theta_2 > \theta_u) = 0.025$$

$$Now,, P(\theta_2 > \theta_u) = 0.025$$

$$we\ can\ say, P\left(\frac{52.5}{10.5}\frac{\theta_2}{1-\theta_2} > \frac{52.5}{10.5}\frac{\theta_u}{1-\theta_u}\right) = 0.025$$

$$\Rightarrow P(F > F_u) = 0.025$$

$$\Rightarrow \frac{52.5}{10.5}\frac{\theta_u}{1-\theta_u} = 1.823$$

$$\Rightarrow \frac{\theta_u}{1-\theta_u} = 0.3646$$

$$\Rightarrow \theta_u = 0.267$$

$$Again, P(\theta_2 \leq \theta_l) = 0.025$$

$$\Rightarrow P\left(\frac{52.5}{10.5}\frac{\theta_2}{1-\theta_2} \leq \frac{52.5}{10.5}\frac{\theta_l}{1-\theta_l}\right) = 0.025$$

$$\Rightarrow P(F < F_l) = 0.025$$

$$\Rightarrow \frac{52.5}{10.5}\frac{\theta_l}{1-\theta_l} = 0.321$$

$$\Rightarrow \frac{\theta_l}{1-\theta_l} = 0.0642$$

$$\Rightarrow \theta_l = 0.0608$$

Hence the credible interval for $\theta_2|X$ *is 0.267 and 0.0608.*

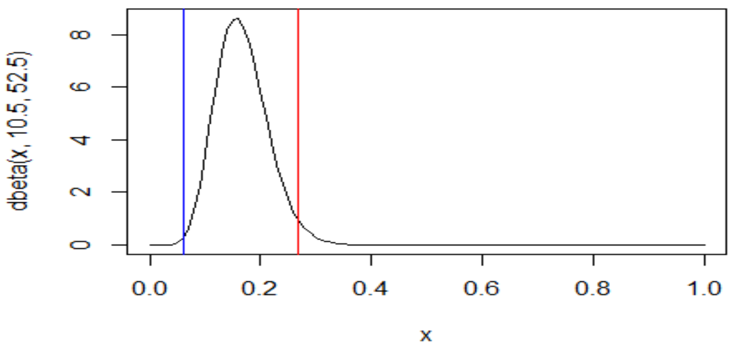

Diagram 02: Graphical representation for Credible Interval of $\theta_2|X$

Credible interval for $\theta_3|X$

$\theta_3|X \sim Beta(x_3 + \alpha_3, N + \alpha - x_3 - \alpha_3)$

$\theta_3|X \sim Beta(12.5, 50.5)$

$\frac{50.5}{12.5}\frac{\theta_3}{1-\theta_3} \sim F_{25,101}$

$Now, P(\theta_l \leq \theta_3 < \theta_u) = 0.95$

$\therefore P(\theta_3 \leq \theta_l) = 0.025$

$and, P(\theta_3 > \theta_u) = 0.025$

$Now,, P(\theta_3 > \theta_u) = 0.025$

$we\ can\ say, P\left(\frac{50.5}{12.5}\frac{\theta_3}{1-\theta_3} > \frac{50.5}{12.5}\frac{\theta_u}{1-\theta_u}\right) = 0.025$

$\Rightarrow P(F > F_u) = 0.025$

$\Rightarrow \frac{50.5}{12.5}\frac{\theta_u}{1-\theta_u} = 1.769$

$\Rightarrow \frac{\theta_u}{1-\theta_u} = 0.437$

$\Rightarrow \theta_u = 0.304$

$Again, P(\theta_3 \leq \theta_l) = 0.025$

$\Rightarrow P\left(\frac{50.5}{12.5}\frac{\theta_3}{1-\theta_3} \leq \frac{50.5}{12.5}\frac{\theta_l}{1-\theta_l}\right) = 0.025$

$\Rightarrow P(F < F_l) = 0.025$

$\Rightarrow \frac{50.5}{12.5}\frac{\theta_l}{1-\theta_l} = 0.501$

$\Rightarrow \frac{\theta_l}{1-\theta_l} = 0.123$

$\Rightarrow \theta_l = 0.109$

Hence the credible interval for $\theta_3|X$ *is 0.304 and 0.109*.

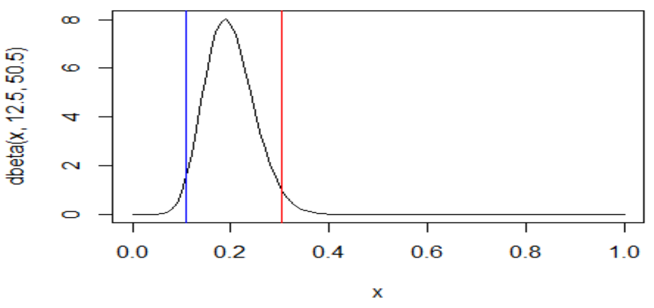

Diagram: 03: Graphical representation for Credible Interval of $\theta_3|X$

Credible interval for $\theta_4|X$

$$\theta_4|X \sim Beta(x_4+\alpha_4, N+\alpha-x_4-\alpha_4)$$

$$\theta_4|X \sim Beta(5.5, 57.5)$$

$$\frac{57.5}{5.5}\frac{\theta_4}{1-\theta_4} \sim F_{11,115}$$

$$Now, P(\theta_l \le \theta_4 < \theta_u) = 0.95$$

$$\therefore P(\theta_4 \le \theta_l) = 0.025$$

$$and, P(\theta_4 > \theta_u) = 0.025$$

$$Now,, P(\theta_4 > \theta_u) = 0.025$$

$$we\ can\ say, P\left(\frac{57.5}{5.5}\frac{\theta_4}{1-\theta_4} > \frac{57.5}{5.5}\frac{\theta_u}{1-\theta_u}\right) = 0.025$$

$$\Rightarrow P(F > F_u) = 0.025$$

$$\Rightarrow \frac{57.5}{5.5}\frac{\theta_u}{1-\theta_u} = 2.106$$

$$\Rightarrow \frac{\theta_u}{1-\theta_u} = 0.201$$

$$\Rightarrow \theta_u = 0.167$$

$$Again, P(\theta_4 \le \theta_l) = 0.025$$

$$\Rightarrow P\left(\frac{57.5}{5.5}\frac{\theta_4}{1-\theta_4} \le \frac{57.5}{5.5}\frac{\theta_l}{1-\theta_l}\right) = 0.025$$

$$\Rightarrow P(F < F_l) = 0.025$$

$$\Rightarrow \frac{57.5}{5.5}\frac{\theta_l}{1-\theta_l} = 0.339$$

$$\Rightarrow \frac{\theta_l}{1-\theta_l} = 0.032$$

$$\Rightarrow \theta_l = 0.031$$

Hence the credible interval for $\theta_4|X$ *is 0.167 and 0.031.*

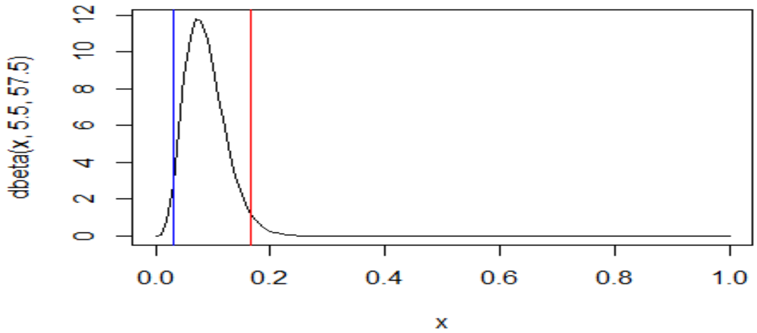

Diagram 04: Graphical representation for Credible Interval of $\theta_4|X$

Credible interval for $\theta_5|X$

$$\theta_5|X \sim Beta(x_5+\alpha_5, N+\alpha-x_5-\alpha_5)$$

$$\theta_5|X \sim Beta(7.5, 55.5)$$

$$\frac{55.5}{7.5}\frac{\theta_5}{1-\theta_5} \sim F_{15,111}$$

$$Now, P(\theta_l \le \theta_5 < \theta_u) = 0.95$$

$$\therefore P(\theta_5 \le \theta_l) = 0.025$$

$$and, P(\theta_5 > \theta_u) = 0.025$$

$$Now,, P(\theta_5 > \theta_u) = 0.025$$

$$we\ can\ say, P\left(\frac{55.5}{7.5}\frac{\theta_5}{1-\theta_5} > \frac{55.5}{7.5}\frac{\theta_u}{1-\theta_u}\right) = 0.025$$

$$\Rightarrow P(F > F_u) = 0.025$$

$$\Rightarrow \frac{55.5}{7.5}\frac{\theta_u}{1-\theta_u} = 1.954$$

$$\Rightarrow \frac{\theta_u}{1-\theta_u} = 0.264$$

$$\Rightarrow \theta_u = 0.208$$

$$Again, P(\theta_5 \le \theta_l) = 0.025$$

$$\Rightarrow P\left(\frac{55.5}{7.5}\frac{\theta_4}{1-\theta_4} \le \frac{55.5}{7.5}\frac{\theta_l}{1-\theta_l}\right) = 0.025$$

$$\Rightarrow P(F < F_l) = 0.025$$

$$\Rightarrow \frac{55.5}{7.5}\frac{\theta_l}{1-\theta_l} = 0.405$$

$$\Rightarrow \frac{\theta_l}{1-\theta_l} = 0.054$$

$$\Rightarrow \theta_l = 0.051$$

Hence the credible interval for $\theta_5|X$ *is 0.208 and 0.051*.

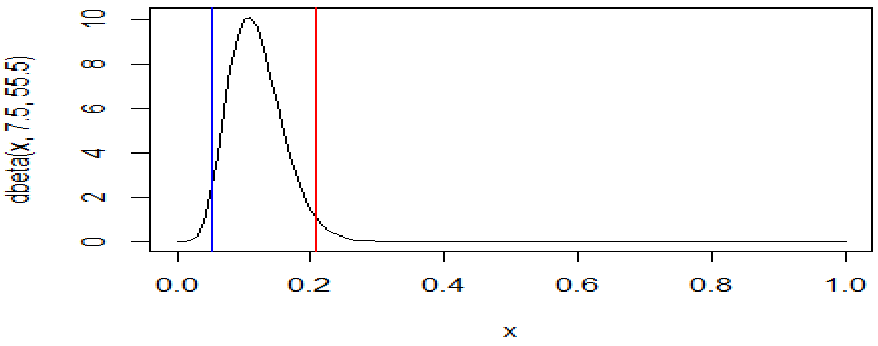

Diagram 05: Graphical representation for Credible Interval of $\theta_5|X$

Credible interval for $\theta_6|X$

$\theta_6|X \sim Beta(x_6 + \alpha_6, N + \alpha - x_6 - \alpha_6)$

$\theta_6|X \sim Beta(18.5, 44.5)$

$\frac{44.5}{18.5}\frac{\theta_6}{1-\theta_6} \sim F_{37,89}$

$Now, P(\theta_l \leq \theta_6 < \theta_u) = 0.95$

$\therefore P(\theta_6 \leq \theta_l) = 0.025$

$and, P(\theta_6 > \theta_u) = 0.025$

$Now,, P(\theta_6 > \theta_u) = 0.025$

$we\ can\ say, P\left(\frac{44.5}{18.5}\frac{\theta_6}{1-\theta_6} > \frac{44.5}{18.5}\frac{\theta_u}{1-\theta_u}\right) = 0.025$

$\Rightarrow P(F > F_u) = 0.025$

$\Rightarrow \frac{44.5}{18.5}\frac{\theta_u}{1-\theta_u} = 1.677$

$\Rightarrow \frac{\theta_u}{1-\theta_u} = 0.697$

$\Rightarrow \theta_u = 0.410$

$Again, P(\theta_6 \leq \theta_l) = 0.025$

$\Rightarrow P\left(\frac{44.5}{18.5}\frac{\theta_6}{1-\theta_6} \leq \frac{44.5}{18.5}\frac{\theta_l}{1-\theta_l}\right) = 0.025$

$\Rightarrow P(F < F_l) = 0.025$

$\Rightarrow \frac{44.5}{18.5}\frac{\theta_l}{1-\theta_l} = 0.559$

$\Rightarrow \frac{\theta_l}{1-\theta_l} = 0.232$

$\Rightarrow \theta_l = 0.188$

Hence the credible interval for $\theta_6|X$ *is 0.410 and 0.188.*

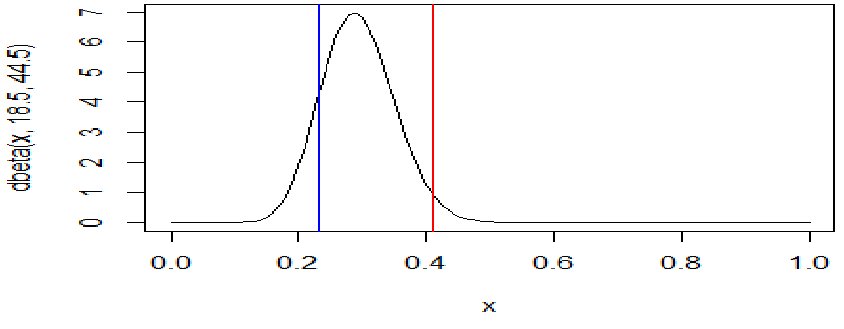

Diagram 06: Graphical representation for Credible Interval for $\theta_6|X$

Credible interval for $\theta_7|X$

$\theta_7|X \sim Beta(x_7 + \alpha_7, N + \alpha - x_7 - \alpha_7)$

$\theta_7|X \sim Beta(3.5, 59.5)$

$\frac{59.5}{3.5}\frac{\theta_7}{1-\theta_7} \sim F_{7,119}$

$Now, P(\theta_l \leq \theta_7 < \theta_u) = 0.95$

$\therefore P(\theta_7 \leq \theta_l) = 0.025$

$and, P(\theta_7 > \theta_u) = 0.025$

$Now,, P(\theta_7 > \theta_u) = 0.025$

$we\ can\ say, P\left(\frac{59.5}{3.5}\frac{\theta_7}{1-\theta_7} > \frac{3.5}{59.5}\frac{\theta_u}{1-\theta_u}\right) = 0.025$

$\Rightarrow P(F > F_u) = 0.025$

$\Rightarrow \frac{59.5}{3.5}\frac{\theta_u}{1-\theta_u} = 2.395$

$\Rightarrow \frac{\theta_u}{1-\theta_u} = 0.14$

$\Rightarrow \theta_u = 0.123$

$Again, P(\theta_7 \leq \theta_l) = 0.025$

$\Rightarrow P\left(\frac{59.5}{5.5}\frac{\theta_7}{1-\theta_7} \leq \frac{59.5}{5.5}\frac{\theta_l}{1-\theta_l}\right) = 0.025$

$\Rightarrow P(F < F_l) = 0.025$

$\Rightarrow \frac{59.5}{3.5}\frac{\theta_l}{1-\theta_l} = 0.238$

$\Rightarrow \frac{\theta_l}{1-\theta_l} = 0.014$

$\Rightarrow \theta_l = 0.0138$

Hence the credible interval for $\theta_7|X$ *is 0.123 and 0.0138.*

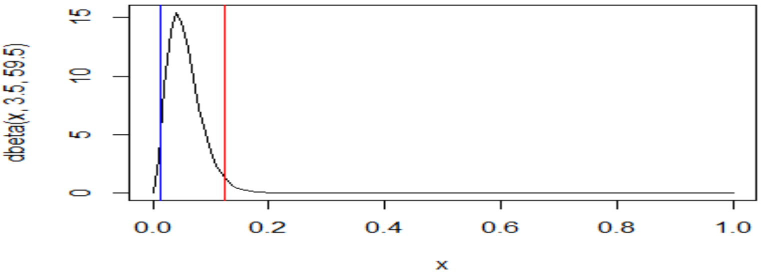

Diagram 07: Graphical representation for Credible Interval for $\theta_7|X$

Credible interval for $\theta_8|X$

$$\theta_8|X \sim Beta(x_8+\alpha_8, N+\alpha-x_8-\alpha_8)$$

$$\theta_8|X \sim Beta(2.5, 60.5)$$

$$\frac{60.5}{2.5}\frac{\theta_8}{1-\theta_8} \sim F_{5,121}$$

$$Now, P(\theta_l \le \theta_8 < \theta_u) = 0.95$$

$$\therefore P(\theta_8 \le \theta_l) = 0.025$$

$$and, P(\theta_8 > \theta_u) = 0.025$$

$$Now,, P(\theta_8 > \theta_u) = 0.025$$

$$we\ can\ say, P\left(\frac{60.5}{2.5}\frac{\theta_8}{1-\theta_8} > \frac{60.5}{2.5}\frac{\theta_u}{1-\theta_u}\right) = 0.025$$

$$\Rightarrow P(F > F_u) = 0.025$$

$$\Rightarrow \frac{60.5}{2.5}\frac{\theta_u}{1-\theta_u} = 2.673$$

$$\Rightarrow \frac{\theta_u}{1-\theta_u} = 0.110$$

$$\Rightarrow \theta_u = 0.099$$

$$Again, P(\theta_8 \le \theta_l) = 0.025$$

$$\Rightarrow P\left(\frac{60.5}{2.5}\frac{\theta_8}{1-\theta_8} \le \frac{60.5}{2.5}\frac{\theta_l}{1-\theta_l}\right) = 0.025$$

$$\Rightarrow P(F < F_l) = 0.025$$

$$\Rightarrow \frac{60.5}{2.5}\frac{\theta_l}{1-\theta_l} = 0.164$$

$$\Rightarrow \frac{\theta_l}{1-\theta_l} = 0.0067$$

$$\Rightarrow \theta_l = 0.0066$$

Hence the credible interval for $\theta_8|X$ *is 0.099 and 0.0066.*

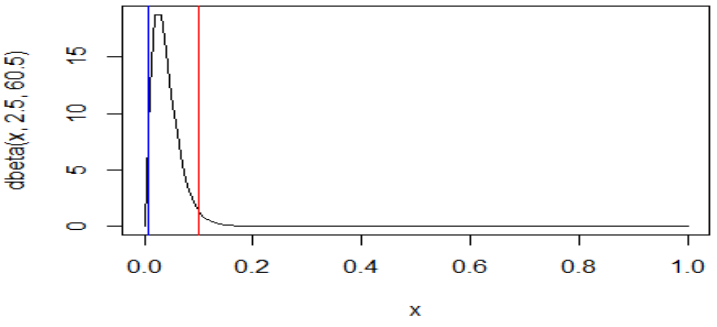

Diagram 08: Graphical representation for Credible Interval for $\theta_8|X$

4.3.1 Data collected

Table 4: A snapshot of the responses for the questionnaire related to the data analysis and '.' Indicates the choice of the responses.

Sample	**Most Worrying Issue in India**							
	Terrorism	Unemployment	Corruption	Environmental Threats	Moral Decline	Religion Extremism	Female Security	Price Hike
F01						•		
F02						•		
F03					•			
F04					•			
F05		•						
F06			•					
F07			•					
F08						•		
F09						•		
F10							•	
F11		•						
F12		•						
F13		•						
F14						•		
F15							•	
F16		•						
F17			•					
F18	•							
F19							•	
F20	•							
.								
.								
.								
F63						•		
Total	4	10	8	2	5	14	18	2
				Grand Total=63				

4.3.2 Data distribution.

In table 02, we see that the total no of observations opted for the mentioned categories (ie. Terrorism, unemployment etc.).

Y_i be the random variable denoting the number of observations voted for the i^{th} choice. (i=1,2,…,8)

So, $(Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8) \sim$ Multinomial $(N, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8)$

The probability mass function will be,

$$f(Y|\phi) = \frac{N!}{y_1!\, y_2!\, y_3!\, y_4!\, y_5!\, y_6!\, y_7!\, y_8!} \times (\lambda_1^{y_1}\ \lambda_2^{y_2}\ \lambda_3^{y_3}\ \lambda_4^{y_4}\ \lambda_5^{y_5}\ \lambda_6^{y_6}\ \lambda_7^{y_7}\ \lambda_8^{y_8})$$

Where, $\sum_{i=1}^{8} x_i = N$, $\sum_{i=1}^{8} \lambda_i = 1$ and $0 \leq \lambda_i \leq 1$, i = 1,2,3,…,8

λ_i's are the parameters of the distribution.

If we put the values of $Y_1, \ldots, Y_8$, the pmf becomes,

$$f(Y|\lambda) = \frac{63!}{4!\, 10!\, 8!\, 2!\, 5!\, 14!\, 18!\, 2!} \times (\lambda_1^{4} \lambda_2^{10} \lambda_3^{8} \lambda_4^{2} \lambda_5^{5} \lambda_6^{14} \lambda_7^{18} \lambda_8^{2}),\ \sum_{i=1}^{8} \lambda_i = 1 \text{ and } 0 \leq \lambda_i \leq 1,\ i = 1,2,3,\ldots,8$$

4.3.3 Construction of prior

To construct the prior distribution, we consider the parameter(s) of the prior distribution as the random variable(s) of the prior distribution. Likewise, in the present scenario, λ_i's will be treated as random variables in the prior distribution. With range, obviously, 0 to 1.

The prior distribution will be,

$(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8) \sim dirichilet(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8)$

The probability density function will be,

$$p(\lambda) = \frac{\Gamma(\alpha)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)\Gamma(\alpha_4)\Gamma(\alpha_5)\Gamma(\alpha_6)\Gamma(\alpha_7)\Gamma(\alpha_8)} \times \quad \lambda_1^{\alpha_1-1} \ldots\ldots \lambda_8^{\alpha_8-1}, \ \alpha_i \geq 0, \ 0 \leq \lambda_i \leq 1$$

4.3.4 Determining posterior.

Hence, the posterior will be,

$\Pi(\lambda|y) \ \alpha \ p(\lambda) \times f(y|\lambda)$

Hence,

$\lambda|Y \sim diri(\alpha_1+y_1, \alpha_2+y_2, \alpha_3+y_3, \alpha_4+y_4, \alpha_5+y_5, \alpha_6+y_6, \alpha_7+y_7, \alpha_8+y_8)$

Where, by Jeffry's prior, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = ½$

If $(Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8)$ ~ Multinomial $(63, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8)$

Then, $\lambda \sim$ diri (1/2, 1/2, 1/2, 1/2, 1/2, 1/2, 1/2, 1/2)

And, $\lambda|Y \sim$ diri(4.5, 10.5, 8.5, 2.5, 5.5, 14.5, 18.5, 2.5)

Jeffrys prior for λ is,

$\alpha_i = 1/2$,i=1,2,3,…,8

4.3.5 Summary.

- Calculation of the marginal from a Dirichlet distribution.

  The integration is similar to 3.4. So, we directly jump to the outcome

$$(\lambda_1|Y) \sim Beta(y_1 + \alpha_1, N + \alpha - y_1 - \alpha_1)$$

i.e. $(\lambda_1|Y) \sim$ Beta (4.5,62.5) (putting the values of α_i's and y_i's)

• Calculation of posterior mean and variance and credible interval.

Table 5: The following table exhibits the posterior mean variance and the credible intervals of the marginal distributions of λ_i's based on the female responses.

Random Variable	Expectation	Variance	Credible interval
$(\lambda_1\|Y) \sim$ Beta(4.5,62.5)	$E(\lambda_1\|Y)$=0.067	$V(\lambda_1\|Y)$=0.00092	(0.0207.0.138)
$(\lambda_2\|Y)\sim$ Beta(10.5,56.5)	$E(\lambda_2\|Y)$=0.158	$V(\lambda_2\|Y)$=0.0019	(0.088,0.337)
$(\lambda_3\|Y) \sim$ Beta(8.5,58.5)	$E(\lambda_3\|Y)$=0.127	$V(\lambda_3\|Y)$=0.00163	(0.0978,0.216)
$(\lambda_4\|Y) \sim$ Beta(2.5,64.5)	$E(\lambda_4\|Y)$=0.037	$V(\lambda_4\|Y)$=0.00053	(0.0064,0.938)
$(\lambda_5\|Y) \sim$ Beta(5.5,61.5)	$E(\lambda_5\|Y)$=0.082	$V(\lambda_5\|Y)$=0.0011	(0.029,0.158)
$(\lambda_6\|Y) \sim$ Beta(14.5,52.5)	$E(\lambda_6\|Y)$=0.216	$V(\lambda_6\|Y)$=0.0025	(0.127,0.322)
$(\lambda_7\|Y) \sim$ Beta(18.5,48.5)	$E(\lambda_7\|Y)$=0.276	$V(\lambda_7\|Y)$=0.0029	(0.176,0.388)
$(\lambda_8\|Y) \sim$ Beta(2.5,64.5)	$E(\lambda_8\|Y)$=0.037	$V(\lambda_8\|Y)$=0.00053	(0.0064,0.938)

Note: The derivation and the diagrammatic representation of the posterior marginal distributions are similar to the derivations shown while analysing the male responses. However, one can find the necessary R codes and the diagrammatic representation in the web appendix.

4.4 Comparative analysis with respect to genders

Now, as per the plan, we wish to conduct a comparative analysis between the male and females based on one single issue. The problem can be redefined as, $P\left(\frac{\theta_7}{\lambda_7} \leq 1\right)$, then to find this type of probabilities, we need to check the probability distribution of the ratio of two independent beta distribution. As we have come across, chapter 3 consists of the detailed analysis of the male population and the females, separately. But the similarity is that in both cases the analysis has been done on the same theme. On that note, there is scope to conduct a comparative analysis. We are keen to do so. The agenda is to investigate (or predict) whether a particular issue (like price hike, terrorism, female security etc.) is being more emphasized by a particular gender and if so, which issue is more important to which gender. There are 8 issues mentioned in the dataset. For example, we are opting two- Religious Extremism and Female security. However, the reader may choose different ones. The necessary calculations remain same. In order to investigate on Female Security, we need to determine whether $P(\theta_7 \geq \lambda_7) < P(\theta_7 < \lambda_7)$ or $P(\theta_7 \geq \lambda_7) > P(\theta_7 < \lambda_7)$ or they are equal. The results are as follows.

$P(\theta_7 \geq \lambda_7)$= 0.0001744 and $P(\theta_7 < \lambda_7) = 0.99982$

Hence, it can be concluded that females are keener to choose "female security" than men are.

Again, let's calculate $P(\theta_6 \geq \lambda_6)$

$P(\theta_6 \geq \lambda_6)$= 0.85

$P(\theta_6 < \lambda_6) = 0.15$

Hence, it can be concluded that males are keener to choose "Religious extremism" than women are.

The mathematical derivations are available in Chapter 3.

The calculation of the probabilities has been done with R- language. The necessary codes are provided in the web appendix section.

5. Discussion and Conclusion

5.1 An overall discussion

This chapter discusses on an overall synopsis of the previous four chapter and thus looks for the relevant conclusion that can be drawn. So far up to chapter 4, the problem is framed and plausible analysis has been carried out in the light of Bayesian methodologies. We are not entering into the debate of philosophical correctness Bayesian methods and frequentist methods but it ought to be mentioned that assumption of priors plays a crucial role in differentiating Bayesian from frequentist methods. Here, in this problem also, starting with a multinomial problem, the prior distribution was set as a Dirichlet distribution which was same for the posterior distribution also (though, the parameters differ). Hence, we conclude the prior as a conjugate prior.

Now, marginal distributions of the random variables were derived and we see that the marginal distributions are nothing but beta distribution of 1^{st} kind. Summary measures were calculated manually. Sometimes use of R language is also introduced. We see that the entire problem deals with two phases. One, the data is based on the male responses and two, the other data is based on female responses. So, based on the same 8 categories, two tables have been formed which necessarily differs with respect to the genders. Considering independence among the responses, we go for the method of ratio of beta random variables. The objective of this task was to determine the proneness of a gender towards a specific issue. This paper discusses explicitly on the application of ratio of two beta random variables in the domain of Bayesian data analysis. Bayesian methods enables us to unfold the data into several dimensions and this effort was one of those many. Besides, the results are obtained using general R functions and simulation technique (Monte Carlo) both. The results obtained from Monte Carlo* simulation is provided below.

Table 05: Posterior mean, variance and credible intervals obtained from Monte Carlo simulation corresponding to the male samples mentioned in chapter 4

Expectation	Variances	Confidence Intervals	
		LL	UL
0.0396	0.0006	0.0067	0.0992
0.1666	0.0022	0.0861	0.2667
0.1986	0.0025	0.1107	0.3045
0.0872	0.0012	0.0312	0.1674
0.1193	0.0016	0.0521	0.2082
0.2936	0.0032	0.1890	0.4109
0.0555	0.0008	0.0137	0.1236
0.0397	0.0006	0.0066	0.0994

* The code is provided in the web appendix section.

Table 06: Posterior mean, variance and credible intervals obtained from Monte Carlo simulation correspond to the male samples mentioned in chapter 4

Expectation	Variances	Confidence Intervals	
		LL	UL
0.0672	0.0009	0.0209	0.1383
0.1568	0.0019	0.0806	0.2529
0.1268	0.0016	0.0588	0.2164
0.0372	0.0005	0.0064	0.0936
0.0820	0.0011	0.0293	0.1581
0.2165	0.0025	0.1277	0.3212
0.2762	0.0029	0.1767	0.3876
0.0373	0.0005	0.0063	0.0939

5.2 Conclusion

The marginal distribution of the parameters follows beta distribution. Based on the discussion on ratio between a pair of independent beta random variables by Pham-Jia (2000), comparative analysis of two pairs of variables, viz., θ_7, λ_7 and θ_6, λ_6 have been conducted. Though further analysis can be conducted within the other variables, we confine our analysis within these two pairs. Form the case 1, (i.e., comparison between θ_7 and λ_7) we observe that the probability of θ_7 being less than λ_7 is sufficiently large. Thus, we can conclude that, while focussing on female security, females are more concerned about it than men.

On the other hand, if we investigate the case 2, we see, the probability of θ_6 being greater than λ_6 is greater than the reverse scenario. Which implies, according to the respondents' mind set, religious extremism is more burning issue to men than to women.

Again, in the 4th segment, whatever has been done in the closed form, is repeated with the Monte Carlo integration method. We have the scope to conclude that, both the processes are applicable under certain constraints. Those constraints are the limitations, we can say. The above illustration is not applicable when the dataset is significantly large. Second, assumption of independent sampling has been considered here. If the samples collected from the two sets of populations are not independent, the application of the concept introduced by Pham-Jia cannot be executed. Finally, it can be admitted that application of Bayesian methods in the inferential statistics enables us to get more information from available dataset(s).

References.

1. O'Hagan, A. (2006). Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering & System Safety*, *91*(10-11), 1290-1300.
2. Pham-Gia, T. (2000). Distributions of the ratios of independent beta variables and applications. *Communications in Statistics-Theory and Methods*, *29*(12), 2693-2715.
3. Rubin, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *The Annals of Statistics*, *12*(4), 1151-1172
4. Berger, J. (2006). The case for objective Bayesian analysis. *Bayesian analysis*, 1(3), 385-402.
5. Raftrey, A. E., & Lewis, S. M. (1996). Implementing mcmc. *Markov chain Monte Carloin* practice,115-130.
6. Gelman, A., Carlin, J. B., Stern, H. S.,Dunson, D. B., Vehtari, A., & Ruin, D.B., (2014). *Bayesian data analysis* (Vol. 2). Boca Raton, FL: CRC press.
7. Lee, P. M. (2012). *Bayesian statistics: an introduction*. John Wiley & Sons.
8. Robert, C.(2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Spring Science &Business Media.
9. Tuyl, F., Gerlach, R., & Mergersen, K.(2008). A comparison of Bayes- Laplace, Jeffreys, and other priors: the case of zero events. The *American Statistician*, 62(1), 40-44.
10. Murphy, K. P. (2007). Conjugate Bayesian analysis of the Gaussian distribution. *def*, *1*(2σ2), 16.
11. Minka, T. (2003). Bayesian inference, entropy, and the multinomial distribution. *Online tutorial*.
12. Thomas, D. L., Iianuzzi, C., & Barry, R. P. (2004). A Bayesian multinomial model for analyzing categorical habitat selection data. *Journal of Agricultural, Biological, and Environmental Statistics*, *9*(4), 432-442.
13. Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological review*, *114*(2), 245.
14. Agresti, A., & Hartzel, J. Analysis: Strategies for Comparing Treatments on a Binary Response with Multi-Centre Data. *Tutorials in Biostatistics: Statistical Methods in Clinical Studies, Volume 1*, 397-421.
15. Tuyl, F., Gerlach, R., & Mengersen, K. (2009). Posterior predictive arguments in favor of the Bayes-Laplace prior as the consensus prior for binomial and multinomial parameters. *Bayesian analysis*, *4*(1), 151-158.
16. Ng, K. W., Tian, G. L., & Tang, M. L. (2011). *Dirichlet and related distributions: Theory, methods and applications* (Vol. 888). John Wiley & Sons.
17. Gelman, A. (2008). Teaching Bayes to Graduate Students in Political Science, Sociology, Public Health, Education, Economics,…. *The American Statistician*, *62*(3), 202-205.
18. Bolstad, W. M., & Curran, J. M. (2016). *Introduction to Bayesian statistics*. John Wiley & Sons.
19. Robert, C. P. (2004). *Monte carlo methods*. John Wiley & Sons, Ltd

****The appendix and necessary R codes can be found in the following web address.***