ISSN 2320-5407 International Journal of Advanced Research (2015), Volume 3, Issue 4, 56-65

ISSN NO. 2320-5407

Journal homepage: http://www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH

RESEARCH ARTICLE

OBJECTIVE VERSUS SUBJECTIVE BAYESIAN INFERENCE: A COMPARATIVE STUDY.

Emmanuel TORSEN
Department of Statistics and Operations Research, Modibbo Adama University of Technology Yola, Adamawa – Nigeria.
e-mail: torsen@mautech.edu.ng

Manuscript Info

Manuscript History:

Received: 15 February 2015
Final Accepted: 22 March 2015
Published Online: April 2015

Key words:

Subjective, Objective, prio information, uniform prior, Jeffreys prior, posterior distribution.

***Corresponding Author**

Emmanuel TORSEN

Abstract

In Bayesian statistics, the choice of the prior distribution is often controversial. Different rules for selecting priors have been suggested in the literature, this is broadly classified into objective (non-informative) and subjective (informative) priors. A fundamental feature of the Bayesian approach to statistics is the use of prior information in addition to the (sample) data. A (proper) subjective Bayesian analysis will always incorporate genuine prior information that genuinely represents prior beliefs, which will help to strengthen inferences about the true value of the parameter and ensure that relevant information about it is not wasted. The (improper) objective Bayesian analysis is not able to do that, since the non-informative prior adds nothing to the likelihood. Data on Diabetic cases (Biomedical Laboratory Medical School University of Verona, Italy) was used for illustration.

INTRODUCTION

Broadly speaking, there are two views on Bayesian probability that interpret the 'probability' concept in different ways. For **objectivists**, probability objectively measures the plausibility of propositions, i.e. the probability of a proposition corresponds to a reasonable belief everyone sharing the same knowledge should share in accordance with the rules of Bayesian statistics, which can be justified by requirements of rationality and consistency (Jaynes, 1986 and Cox, 2001). Requirements of rationality and coherence are important for **subjectivists**, for which the probability corresponds to a 'personal belief' (de Finetti, 1974). For subjectivists however, rationality and coherence constrain the probabilities a subject may have, but allow for substantial variation within those constraints. The objective and subjective variants of Bayesian probability differ mainly in their interpretation and construction of the prior probability.

The paper is aimed at making a comparative study between the use (or attitude to use) of informative and non-informative priors, digging into the meaning of the two classes of priors and the different reasons in support of one or the other, and use data on the cases of diabetes in order to illustrate the theoretical points.

PRIOR INFORMATION

A random variable can be thought of as a variable that takes on a set of values with specified probability. In Frequentist statistics, parameters are not repeatable random things but are fixed (albeit unknown) quantities, which means that they cannot be considered as random variables. In contrast, in Bayesian statistics anything about which we are uncertain, including the true value of a parameter, can be thought of as being a random variable to which we can assign a probability distribution, known specifically as **prior information**. A fundamental feature of the Bayesian approach to statistics is the use of prior information in addition to the (sample) data. A proper Bayesian analysis will always incorporate genuine prior information, which will help to strengthen inferences about the true value of the parameter and ensure that any relevant information about it is not wasted.

There are two types of priors: informative and non-informative. Box and Tiao (1973) defined a non-informative prior as one that provides little information relative to the experiment - in this case, the diabetic cases data. Informative prior distributions, on the other hand, summarize the evidence about the parameters concerned from many sources and often have a considerable impact on the results.

INFORMATIVE PRIOR

In general, the argument against the use of prior information is that it is intrinsically subjective and therefore has no place in science. Of particular concern is the fact that an unscrupulous analyst can concoct any desired result by the creative specification of prior distributions for the parameters in the model.

However, the potential for manipulation is not unique to Bayesian statistics. The scientific community (Bolstad, 2004) and regulatory agencies (Gilks and Spiegelhalter, 1996) have developed sophisticated safeguards and guidance to avoid conscious or unconscious biases.

NON-INFORMATIVE PRIOR

To avoid having to use informative prior distributions while still being able to use Bayesian tools, some authors have suggested using non-informative prior distributions to represent a state of prior ignorance (Briggs, 1999). In so doing the analyst obtains some of the benefits of a Bayesian approach, particularly that the results are presented in the intuitive way in which one would like to make inferences. Another way to represent prior information is to specify skeptical prior distributions for the parameters in the model. In this context, the prior distribution is specified in such a way that it automatically favors the standard treatment. Such a proposal is tempting, particularly in a regulatory framework where rigorous standards and safeguards are demanded. However, both ideas suffer from serious objections, and both fail to exploit the full potential of the Bayesian approach. In addition, there is no unique way to implement either idea, and hence subjectivity is not removed.

THE BAYESIAN APPROACH

A Bayesian analysis synthesizes two sources of information about the unknown parameters of interest. The first of these is the sample data, expressed formally by the **likelihood function**. The second is the prior distribution, which represents additional (external) information that is available to the investigator (Figure 1). Whereas the likelihood function is also fundamental to Frequentist inference, the prior distribution is used only in the Bayesian approach.

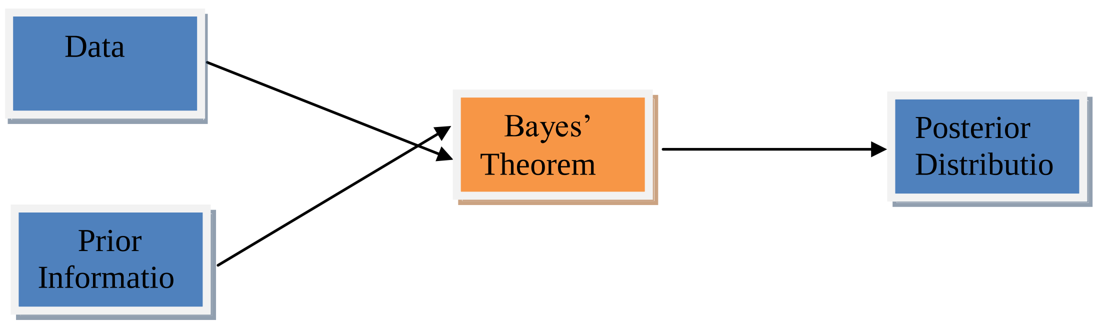


Figure 1: The Bayes Method

If we represent the data by the symbol X and denote the set of unknown parameters by Θ, then the likelihood function is $L(\theta/x)$; the probability of observing the data X being conditional on the values of the parameter θ. If we further represent the prior distribution for θ as $\pi(\theta)$, giving the probability that θ takes any particular value based on whatever additional information might be available to the investigator, then, with the application of Bayes' theorem, (Senn, 2003) an elementary result about conditional probability named after the Reverend Thomas Bayes, we synthesize these two sources of information through the equation:

$$P(\theta/x) \propto \pi(\theta)L(\theta/x) \qquad (1)$$

The proportionality symbol $\propto$ expresses the fact that the product of the likelihood function and the prior distribution on the right hand side of Equation (1) must be scaled to integrate to one over the range of plausible θ values for it to be a proper probability distribution. The scaled product $P(\theta/x)$, is then called the **posterior** distribution for θ (given the data), and expresses what is now known about θ based on both the sample data and prior information.

BAYESIAN INFERENCE

The Bayesian statistical inference can be approached from two different angles: from the Subjectivist point of view and from the Frequentist (Objectivist) point of view.

The Subjective approach is the one by which the prior probability density function is defined as a subjective opinion of the person involved in the inferential process, while the Frequentist approach (sometimes referred to as the Objective approach) is the one by which the prior density function is defined in terms of some empirical evidence only. This is the main concern of this study.

OBJECTIVE BAYESIAN APPROACH

Usually, the Researcher supporting this type of approach would like to work with non-informative prior in order not to change the information brought by the likelihood. Consequently, to make the posterior probability density function using only the information brought by the likelihood.

In the vast literature concerned with the problem of a non-informative prior, there are three (3) major proposals. The easiest one to use is the Uniform prior probability density function and then the Harold Jeffrey's Prior and more recently the Reference prior (Berger-Bernardo).

UNIFORM PRIOR

About 200 years ago, both Bayes (1763) and Laplace (1814) postulated that: "When nothing is known about θ in advance, let the prior π(θ), be a uniform distribution, that is, let all possible outcomes of θ have the same probability", and this is also known as the "principle of insufficient reason", Laplace (1814).

R. A. Fisher (similar to the Subjectivist) did not support the Bayes/Laplace postulate; he argued that "not knowing the chance of mutually exclusive events and knowing the chance to be equal are two quite different states of knowledge". He accepted Bayes' theorem only for informative priors (Syversveen, 2003).

Laplace (1812), judged that, it worked exceptionally well to simply always choose the prior for θ to be the constant $\{\pi(\theta) = k\}$ on the parameter space Θ.

The Uniform prior (non-informative for) distribution of θ is:

$$\pi(\theta) = k = \frac{1}{(b-a)}, \int_a^b \pi(\theta) d\theta = 1; \quad \theta \in \Theta = I_{(a;b)} \tag{2}$$

The posterior probability density function of parameter β obtained by a Uniform prior probability density function $\pi(\beta)$ and a Likelihood function for a given data vector $\tilde{x}_n = \{x_1, \dots, x_n\}$ represented by a Gamma function, is:

$$P(\beta/x) \propto \pi(\beta) L(\alpha, \beta) = 1 \times L(\alpha, \beta) = \mathrm{Gamma}(\alpha, \beta) \tag{3}$$

$$L(\alpha, \beta) \propto T_2^{\alpha-1} e^{-\beta T_1} \tag{4}$$

Where $T_1 = \sum_{i=1}^{n} x_i$ and $T_2 = \prod_{i=1}^{n} x_i$

Which is equal to the density kernel of the Gamma distribution with parameters α and β. Hence, the posterior distribution given data is a Gamma (α,β).

- The uniform prior is a diffused prior.
- The domain of the parameter θ in the experiment (likelihood) is the sub-domain of the prior.

JEFFREY'S PRIOR

NOTE:

- Jeffrey noted the weakness of constant prior: accepting a state of ignorance about the parameter θ for all the parameter space Θ. For example, consider making an inference about the probability θ of getting Head from tossing a coin, we may feel that, the prior probability density function to be uniformly distributed around 0.5, but the prior probability density function should not be the same for values of θ nearby 0 or nearby 1.
- Jeffrey did not want to use and to mix personalistic opinions about the parameters with the information brought by the data (alias: the Likelihood).
- He chose to use the Fisher information because it tells us how much information is in the likelihood about θ. In other words, Fisher information I(θ) is an indicator of the amount of information brought by the model (observations) about θ.
- Hence, he (Jeffrey) developed a prior that is not dependent upon the set of parameter variables that is chosen to describe parameter space Θ, this is to favor the values for θ of which I(θ) is large is equivalent to minimizing the influence of the prior.

Jeffrey's prior distribution is defined as the density of the parameter proportional to the square root of the determinant of the Fisher Information, is given by:

$$\pi(\theta) \propto \sqrt{I_n(\theta)} \tag{5}$$

Where $I_n(\theta) = E_\theta \left\{ \left(\frac{\partial \log f(\theta)}{\partial \theta} \right)^2 \right\}$

$$I_n(\theta) = -E_\theta \left\{ \frac{\partial^2 \log_e L(\theta)}{\partial \theta} \right\}$$

THE LIKELIHOOD FUNCTION

The likelihood principle states that, all evidence which is obtained from an experiment, about an unknown quantity θ, is contained in the likelihood function of θ for the given data.

DEFINITION: Let $\tilde{x}_n = \{x_1, \dots, x_n\}$ have joint density $f(\tilde{x}_n; \theta) = f(x_1, \dots, x_n; \theta)$.

The likelihood function $L: \Theta \to [0, \infty)$ is defined by

$$L(\theta) \equiv L(\theta; \tilde{x}_n) = f(\tilde{x}_n; \theta) \tag{6}$$

Where $\tilde{x}_n$ is fixed and θ varies in Θ.

- The likelihood function is a function of θ.
- The likelihood function is not a probability density function.
- If the data are iid, then the likelihood is

$$L(\theta) = \prod_{i=1}^{n} f(\tilde{x}_n; \theta), \text{ iid case only.} \tag{7}$$

- The likehood function is only defined up to a constant of proportionality.
- The likelihood function is used (i) to generate estimators (the maximum likelihood estimator) and (ii) as a key ingredient in Bayesian inference.

Let a random variable X having parameters α and β follow the Gamma distribution. So the probability density function (p.d.f) of the Gamma distribution for α and β is:

$$f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \; x > 0, \; \alpha > 0, \; \beta > 0 \tag{8}$$

The Gamma p.d.f is used here because; the data used as example for this work follows a Gamma p.d.f. The shape parameter (α) is assumed to be known and the scale parameter (β) is unknown.
The likelihood function for an independently and identically distributed (iid) random variable $X_i, i = 1,2,\dots,n$ of size n is given by:

$$L(\alpha,\beta) = \prod_{i=1}^{n} f(x_i|\alpha,\beta)$$

$$L(\alpha,\beta) = \prod_{i=1}^{n}\left[\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}\right] \propto {T_2}^{\alpha-1}e^{-\beta T_1} \quad (9)$$

$$log_e\, L(\alpha,\beta) = n\alpha\, log_e\,\beta - n\, log_e\,\Gamma(\alpha) + (\alpha-1)\sum_{i=1}^{n} log_e\, x_i - \beta T_1 \quad (10)$$

$$\frac{\partial\, log_e\, L(\alpha,\beta)}{\partial\alpha} = n\, log_e\,\beta - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n} log_e\, x_i$$

$$\frac{\partial\, log_e\, L(\alpha,\beta)}{\partial\beta} = \frac{n\alpha}{\beta} - T_1$$

$$\frac{\partial^2\, log_e\, L(\alpha,\beta)}{\partial\beta} = -\frac{n\alpha}{\beta^2} \quad (11)$$

Hence, the Fisher information for β is:

$$I_n(\beta) = -E_\beta\left\{\frac{\partial^2\, log_e\, L(\alpha,\beta)}{\partial\beta}\right\} = \frac{n\alpha}{\beta^2} \quad (12)$$

The likelihood function is:

$$L(\alpha,\beta) \propto {T_2}^{\alpha-1}e^{-\beta T_1} \quad (13)$$

Where $T_1 = \sum_{i=1}^{n} x_i$ and $T_2 = \prod_{i=1}^{n} x_i$

In the situation where one does not have much information about the parameters, Jeffrey (1946), suggested a non-informative prior.
The Jeffrey's prior for β is:

$$\pi(\beta) \propto \sqrt{I_n(\beta)} = \sqrt{\left[-E_\beta\left\{\frac{\partial^2\, log_e\, L(\alpha,\beta)}{\partial\beta}\right\}\right]} = \sqrt{\frac{n\alpha}{\beta^2}} \propto 1 \quad (14)$$

Hence the posterior distribution using (14) and (13) is given by:

$$P(\beta/x) \propto \pi(\beta)L(\alpha,\beta) = 1 \times L(\alpha,\beta) = Gamma(\alpha,\beta) \quad (15)$$

$$P(\beta/x) \propto {T_2}^{\alpha-1}e^{-\beta T_1} \quad (16)$$

Which is the density kernel of the Gamma distribution with parameters α and β. Hence, the posterior distribution given data is Gamma (α,β).

SUBJECTIVE BAYESIAN APPROACH

The two-parameter Gamma distribution has one shape parameter and one scale parameter. The random variable X follows Gamma distribution with the shape and the scale parameters as α >0 and β >0 respectively, if it has the probability density function (p.d.f.) as given in equation (5) above. It will be denoted by $Gamma(\alpha,\beta)$. Hence, $\Gamma(\alpha)$ is the Gamma function and it is expressed as

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx \quad (17)$$

It is well known that the p.d.f. of $Gamma(\alpha,\beta)$ can take different shapes but it is always unimodal. The moments of Xcan be obtained in explicit form as;

$$E(X) = \frac{\alpha}{\beta}, \quad and \quad V(X) = \frac{\alpha}{\beta^2} \quad (18)$$

It is assumed that $\{x_1, x_2, \dots, x_n\}$ is a random sample from $f(./\alpha,\beta)$. It is also assumed that β has a prior $\pi_1(\beta)$, and $\pi_1(\beta)$ follows $Gamma(\alpha_0,\beta_0) = Gamma(2,1)$.

$$\pi_1(\beta) = \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha)}\beta^{\alpha_0-1}e^{-\beta_0\beta}, \alpha_0 > 0, \beta_0 > 0, \beta > 0 \quad (19)$$

At this moment we do not assume any specific prior on α. It is simply assumed that the prior on α is $\pi_2(.)$ and the density function of $\pi_2(.)$ is log-concave and it is independent of $\pi_1(.)$.
The likelihood function of the observed data is

$$L(\alpha,\beta) = \frac{\beta^{n\alpha}}{(\Gamma(\alpha))^n}e^{-\beta T_1}T_2^{\alpha-1}, \quad (20)$$

Where $T_1 = \sum_{i=1}^{n} x_i$ and $T_2 = \prod_{i=1}^{n} x_i$. Note that (T_1, T_2) are jointly sufficient for (α,β).
Therefore, the joint density function of the observed $data$, α and β is

$$L(data;\alpha,\beta) \propto \frac{1}{(\Gamma(\alpha))^n}\beta^{\beta_0+n\alpha-1}\, e^{-\beta(\alpha_0+T_1)}T_2^{\alpha-1}\pi_2(\alpha) \quad (21)$$

The posterior density function of $\{\alpha,\beta\}$ given the data is

$$P(\alpha,\beta;data) = \frac{\frac{1}{(\Gamma(\alpha))^n}\beta^{\alpha_0+n\alpha-1}\, e^{-\beta(\beta_0+T_1)}T_2^{\alpha-1}\pi_2(\alpha)}{\int_0^\infty\int_0^\infty\frac{1}{(\Gamma(\alpha))^n}\beta^{\alpha_0+n\alpha-1}\, e^{-\beta(\beta_0+T_1)}T_2^{\alpha-1}\pi_2(\alpha)d\alpha d\beta}$$

Therefore, posterior density function of β given the data is

$$P(\beta;data) = \frac{1}{(\Gamma(\beta_0+n\alpha))^n}\beta^{\alpha_0+n\alpha-1}\, e^{-\beta(\beta_0+T_1)}T_2^{\alpha_0+n\alpha-1} \quad (22)$$

$$P(\beta;data) \propto \beta^{\alpha_0+n\alpha-1}\, e^{-\beta(\beta_0+T_1)}T_2^{\alpha_0+n\alpha-1} \quad (23)$$

Hence, the posterior density function of β given the data is

$$Gamma(\alpha_0+n\alpha, \beta_0+T_1) = Gamma(\alpha^*, \beta^*). \quad (24)$$

Where $\alpha^* = \alpha_0 + n\alpha$ and $\beta^* = \beta_0 + T_1$

CREDIBLE INTERVAL (PROBABILITY INTERVAL)

If $P(\theta/\tilde{x})$ is the posterior density of θ, then any interval (a, b) satisfying

$$P\{\theta\epsilon(a,b)/\tilde{x}\} = \int_a^b P(\theta/\tilde{x})d\theta = p, \quad (25)$$

Where p is a constant, is called a 100p% Credible Interval.

EQUAL-TAILED INTERVAL

The Equal-tailed 100p% Credible Interval for which a and b are chosen to satisfy the equation;

$$\int_{-\infty}^{a} P(\theta/\tilde{x})d\theta = \int_{b}^{\infty} P(\theta/\tilde{x})d\theta = \frac{1}{2}(1-p) \quad (26)$$

Now, the posterior density here is a Gamma density function. The Gamma posterior distribution can be converted to a Chi-squared distribution since the gamma table values are not readily available.

That is, if $\theta \sim \Gamma(\alpha,\beta) = \theta \sim Gamma(\alpha,\beta)$, then

$$2\beta\theta \sim \chi^2_{2(\alpha)}$$
$$\theta \sim \frac{\chi^2_{2(\alpha)}}{2\beta}$$

Let (a, b) be the 95% Credible Interval for θ. Then,

$$P[a \le \theta \le b] = 0.95$$
$$P\left[a \le \frac{\chi^2_{2(\alpha)}}{2\beta} \le \mathrm{b}\right] = 0.95$$
$$P\left[2\beta\theta a \le \chi^2_{2(\alpha)} \le 2\beta\theta b\right] = 0.95$$
$$2\beta a = \chi^2_{1(0.975,2\alpha)} \quad \text{and } 2\beta b = \chi^2_{2(0.025,2\alpha)}$$
$$a = \frac{\chi^2_{1(0.975,2\alpha)}}{2\beta} \quad \text{and} \quad b = \frac{\chi^2_{2(0.025,2\alpha)}}{2\beta}$$

Thus, the 95% Credible Interval for θ is (a, b). That is, the probability that θ lies between (a, b) is $(1 - p)$.

DATA ANALYSIS

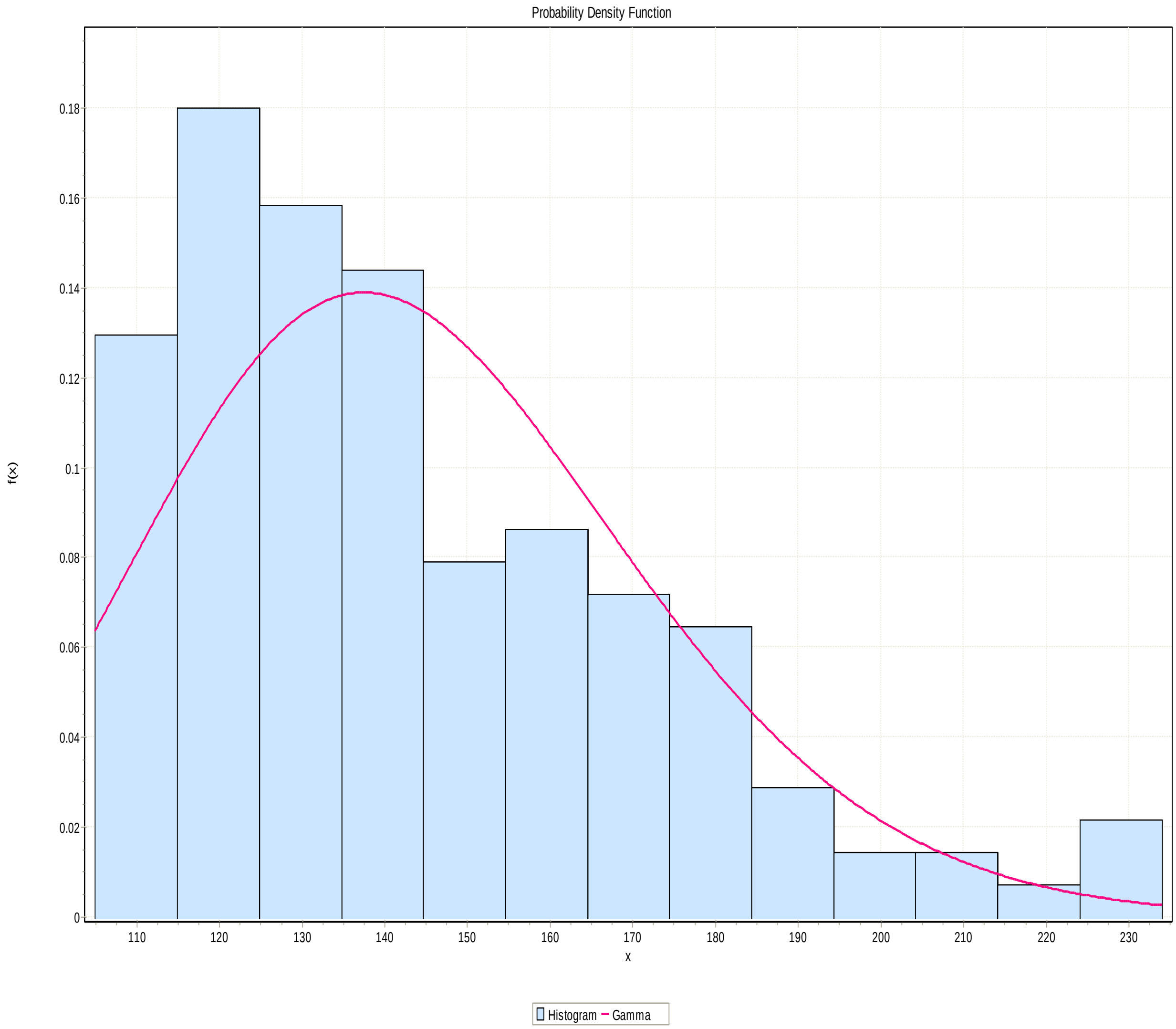
Software used: EasyFit 5.5 and R.

Table1: Descriptive Statistics

Sample Size (n)	139
Range	129
$\sum \boldsymbol{X_i}$	19931
Mean	143.39
Variance	841.83
Standard Deviation	29.014
Coefficient Of Variation	0.20235
Standard Error	2.461

Skewness	1.0425
Excess Kurtosis	0.77583

Figure2: Showing the Histogram (the empirical distribution) and Gamma model (the theoretical distribution)



GOODNESS OF FIT DETAILS

Table2a

lmogorov-Smirnov Test			
Sample Size	139		
Statistic	0.10366		
P-value	0.09389		
α	0.05	0.02	0.01
Critical Value	0.11518	0.12876	0.13817
Reject?	No	No	No

Table2b

Anderson-Darling Test			
Sample Size	139		
Statistic	2.1471		
A	0.05	0.02	0.01
Critical Value	2.5018	3.2892	3.9074
Reject?	No	No	No

Table2c

Chi-Squared Test			
Deg. Of Freedom	7		
Statistic	16.507		
P-value	0.02086		
A	0.05	0.02	0.01
Critical Value	14.067	16.622	18.475
Reject?	Yes	No	No

The tables (2a, 2b and 2c) shows that the data follows a Gamma p.d.f.

THE MODEL PARAMETERS

The model is a Gamma distribution with the following parameters:

$$\alpha = 24.4193, \beta = 0.1703$$

That is;

$$Gamma(\alpha, \beta) = Gamma(24.4193, 0.1703)$$

The Maximum Likelihood Estimate of β is given as;

$$\hat{\beta} = \frac{\alpha}{\bar{X}} = \frac{24.4193}{143.39} = 0.1703$$

SUMMARISING THE OBJECTIVE POSTERIOR DENSITIES

The posterior density is a Gamma probability density function using the two types of objective (Uniform and Jeffrey's) priors considered.

From equation (4) and equation (16), the Gamma posteriors have the following estimates;

$$Posterior\ Mean = \frac{\alpha}{\beta}$$

$$Posterior\ Variance = \frac{\alpha}{\beta^2}$$

$$Posterior\ SD = \sqrt{\frac{\alpha}{\beta^2}}$$

$$Posterior\ Mode = \frac{\alpha - 1}{\beta}$$

Table3

	Posterior Density Estimates	
	Using uniform prior	Using Jeffrey's prior
Posterior Mean	143.39	143.39
Posterior Variance	841.93	841.93
Posterior SD	29.014	29.014
Posterior Mode	137.52	137.52

SUMMARISING THE SUBJECTIVE POSTERIOR DENSITY

The posterior density is a Gamma probability density function using the subjective priors considered.
From equation (24), the Gamma posterior p.d.f., $Gamma(\alpha^*, \beta^*)$ have the following estimates;

$$Posterior\ Mean = \frac{\alpha^*}{\beta^*}$$

$$Posterior\ Variance = \frac{\alpha^*}{(\beta^*)^2}$$

$$Posterior\ S.D = \sqrt{\frac{\alpha^*}{(\beta^*)^2}}$$

$$Posterior\ Mode = \frac{\alpha^*-1}{\beta^*}$$

Table4

Posterior Mean	0.1704
Posterior Variance	$8.5487\ X\ 10^{-6}$
Posterior SD	$2.9238\ X\ 10^{-3}$
Posterior Mode	0.1703

COMPARISON BASED ON BAYESIAN POINT ESTIMATES (MEAN AND MODE)

The Bayesian Point estimates are presented below. The posterior mean and mode using the two approaches are clearly different. Hence, the subjective approach have a better performance because it estimate described our parameter β of interest better.

Table5

	Bayes Estimate	
	Objective Approach	143.39
Posterior Mean	Subjective Approach	0.1704
Posterior Mode	Objective Approach	137.52
	Subjective Approach	0.1703

COMPARISON OF THE TWO APPROACHES (POSTERIOR PDFs) USING THEIR CREDIBLE INTERVALS.

i) Objective Approach:

The two pdfs have the same parameters, hence
The 95% Credible interval for the posterior mean:

```
qgamma(.025, 24.4193, 0.1703)
[1] 92.26539
> qgamma(.975, 24.4193, 0.1703)
[1] 205.6045
```

Thus, the 95% Credible interval is (92.26539, 205.6045). This means that, there is a 95% chance that the mean is in the interval (92.26539, 205.6045).

ii) Subjective Approach:

The 95% Credible interval for the posterior mean:

```
> qgamma(0.025, 3396.2827, 19932)
[1] 0.1647106
> qgamma(0.975, 3396.2827, 19932)
[1] 0.1761714
```

Thus, the 95% Credible interval is (0.1647, 0.1762). This means that, there is a 95% chance that the mean is in the interval (0.1647, 0.1762).

Again it can be seen above that, the Bayesian Credible interval using the subjective approach describes (or is about) the parameter of interest better as compared it objective counterpart.

COMPARISON OF THE TWO APPROACHES (POSTERIOR PDFs) USING COEFFICIENT OF SKEWNESS (COS)

$$COS = \gamma = 2\sqrt{\frac{1}{\alpha}} \tag{27}$$

Table6

		Posterior parameter	Coefficient Of Skewness
		(α, β)	γ
The Approach	Objective	(24.4193, 0.1703)	0.40470
	Subjective	(3396.2827, 19932)	0.03432

From table6, it is observed that $\gamma > 0$ therefore, the posterior distributions based on the Objective Approach are not symmetrical; rather they are both slightly positively and equally skewed. The posterior distribution based on the Subjective Approach approximately symmetrical, this ensures it better performance.

COMPARISON OF THE TWO APPROACHES (POSTERIOR PDFs) WITH RESPECT TO THEIR POSTERIOR VARIANCES.

Table7

Variance Using	
Objective Approach	Subjective Approach
841.93	$8.5487\ X\ 10^{-6}$

From the table7 above, it is obvious that using the Objective Approach, the posterior variance (841.93) is high; this implies low (small) precision. But with the Subjective Approach, the posterior variance (8.5487 X 10^{-6}) is very small this implies very high precision. This means that, the posterior density function which uses the Subjective Approach have much more information about the parameter β of interest. Hence, the Subjective Approach is more efficient.

SUMMARY

In this study, the relative performance of the two approaches to Bayesian Inference (Objective and Subjective) was examined using different performance measures. The comparison of the two approaches is based on the Bayes point estimates, posterior variance, Bayesian credible interval and the coefficient of skewness of the posterior distribution.

CONCLUSION

It has been found that the relative performance of the subjective approach (informative prior) is more efficient than its objective (non-informative priors) counterpart using different performance measures. In fact, it was observed that, while the posterior probability density function obtained using the subjective approach effectively described the unknown parameter β, the posterior probability density function obtained using the objective approach failed to do that.

REFERENCES

- Bernardo, J (1979). *Reference posterior distributions for Bayesian inference*. J.R. Stat. Soc.
- Bernardo, J. and A. Smith (1994). *Bayesian Theory*. John Wiley & Sons.
- Bernardo, J. M. (2005). *Reference Analysis. Handbook of Statistics*. 25 (D. K. Dey and C. R. Rao eds). Amsterdam: Elsevier.
- Bernardo, J. M.(2006). *A Bayesian Mathematical Statistics Primer*.
- Bishop, C.M. (2007). *Pattern Recognition and Machine Learning*. Springer
- Bolstad W.M. (2004). *Introduction to Bayesian Statistics*. Wiley & Sons, Inc., New York
- Box, G. and G. Tiao (1973). *Bayesian Inference in Statistical Analysis*. John Wiley & Sons.
- Cox, D.R. (2006). *Principles of Statistical Inference*. Cambridge University Press.
- Cox, Richard T. (2001). *Algebra of Probable Inference*. The Johns Hopkins University Press.
- Fienberg, Stephen. E. (2006). *When did Bayesian Inference become "Bayesian"? Bayesian Analysis*, 1
- Jaynes, E.T. (1986). *"Bayesian Methods: General Background." In Maximum-Entropy and Bayesian Methods in Applied Statistics, by J. H. Justice (ed.). Cambridge:* Cambridge Univ. Press.
- Jeffreys, H (1946). *An invariant form for the prior probability in estimation problems*. Proc. Roy. Soc.
- Kass, R. and L. Wasserman (1996). *The selection of prior distributions by formal rules*. Journal of the American Statistical Association 91(431).
- Laplace, P. S. (1814). *English edition 1951, A Philosophical Essay on Probabilities*. New York: Dover Publications Inc.
- Pericchi, L. and P. Walley (1991). *Robust Bayesian Credible Intervals and Prior Ignorance*. Int. Statist. Rev. 58(1).
- Pradhan B. and Kundu D. (2012). *Bayes Estimation and Prediction of the Two-Parameter Gamma Distribution*.
- Syversveen A.R. (2003). *Non-informative Bayesian Priors: Interpretation and Problems with Construction and Application*.
- Tahir M. and Hussain Z. *Comparison of Non-informative Priors for Number of Defects (Poisson) Model*.
- Tibshirani, R. (1989). *Noninformative priors for one parameter of many*. Biometrika 76(3).
- Yu J. (2004). *Statistical Decision Theory: Prior Information and Subjective Probability*. Peking University, Beijing.
- Zellner, A. (1996). *An Introduction to Bayesian Inference in Econometrics*. New York; Chichester: John Wiley.
- Zhu M. and Lu A.Y. (2004). *The Counter-intuitive Non-informative prior for the Bernoulli Family*. Journal of Statistical Education, Volume 12, Number2.