We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

5,300
Open access books available

130,000
International authors and editors

155M
Downloads

154
Countries delivered to

Our authors are among the
TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX CLARIVATE ANALYTICS INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Bayesian Inference Application

Wiyada Kumam, Plern Saipara and Poom Kumam

Additional information is available at the end of the chapter

Abstract

In this chapter, we were introduced the concept of Bayesian inference and application to the real world problems such as game theory (Bayesian Game) etc. This chapter was organized as follows. In Sections 2 and 3, we present Model-based Bayesian inference and the components of Bayesian inference, respectively. The last section contains some applications of Bayesian inference.

Keywords: statistical inference, Frequentist inference, Bayesian inference

1. Introduction

In statistical inference, there are two ways for interpretations of probability include Frequentist (or Classical) inference and Bayesian inference. It usually is unlike with each other in the classical nature of probability. Classical inference defines probability as the limit of an event's relative frequency for a large number of experiments and only in the sense of random experiments which are well defined. Other side, Bayesian inference can to impose probabilities to each statement when a random process is not associated. In the sense of Bayesian, probability is a way to show an individual's degree of believes in a statement. Bayesian inferences are different interpretations of probability, and also different approaches depend on those interpretations. Bayes' theorem presents the relativity about two conditional probabilities that are the reverse of anything other. The initials of the term Bayes' theorem is in honor of Reverend Thomas Bayes, and is referred to as Bayes' law (see [1]). This theorem shows the conditional probability or *posterior probability* of an event A after B is observed in terms of the *prior probability* of A, prior probability of B and the conditional probability of B given A. It is valid in all interpretations of probability. Bayes' formula is how to revise probability statements using data. The Bayes' law (or Bayes' rule) is

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{1}$$

The conditional probability definition is defined as follows

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A). \tag{2}$$

For example, let a dice is thrown under a dice-box. From the standard model, all of outcomes have probability equal to 1/6. Now, the dice is lifted a bit and a random corner of the upper side is able to visible which it contains a dot. The new probability distribution of the outcomes shows as follows. Let A_i is the outcome of the throw, for $i = 1, 2, 3, 4, 5, 6$ and B is the randomly chosen corner contains a dot. So, we get $P(A_i) = 1/6$ and $P(B) = 2/3$. We get the following table:

A_i	$P(A_i)$	$P(B \mid A_i)$	$P(A_i \cap B)$	$P(A_i \mid B)$
A_1	1/6	0	0	0
A_2	1/6	1/2	1/12	1/8
A_3	1/6	1/2	1/12	1/8
A_4	1/6	1	1/6	1/4
A_5	1/6	1	1/6	1/4
A_6	1/6	1	1/6	1/4

The simplest way to construct the fourth column is to multiply. For any A_i, $P(A_i | B)$ and $P(B | A_i)$, to sum these values and divide by this sum. This final term is said to be scaling and corresponds to the formula as

$$\sum_{i=1}^{6} P(B|A_i)P(A_i) = \sum_{i=1}^{6} P(A_i \cap B) = P(B).$$

An simpler argument is that $P(A_i | B)$ has to be a probability distribution, thus sum to unity. As the scaling operation is trivial, Bayes' rule is also shown as

$$P(A|B) \propto P(A)P(B|A)$$

where $P(A)$ the prior (distribution), $P(B|A)$ is the likelihood and $P(A|B)$ is the posterior (distribution).

The main result of Bayesians statistics is that statistical inference may depend on the simple device *posterior* $\propto$ *prior* $*$ *likelihood*. By dice-throwing example is not of controversial. The disputations about the possibility of using Bay's rule as

$$P(Truth|Data) = \frac{P(Data|Truth)P(Truth)}{P(Data)}. \tag{3}$$

So, we get

$$P(Truth) = the\ prior. \tag{4}$$

The second ingredient we need is data, plus a how the data associate to the truth which is nothing but the classical concept of specifying a random relationship

$$P(Data|Truth) = the\ likelihood \tag{5}$$

for all associated values of *Truth*. Note that $P(Data|Truth)$ is not applied as probability distribution for different data, but as the probability of the given data for different values of *Truth*. Various authors do apply $P(Data|Truth)$ for likelihood to sheer this misconstrue.

Now, noting that (replace *Truth* with *T*), probability of Data ($P(Data)$) can be written as

$$P(Data) = \int P(T)P(Data|T)dT \tag{6}$$

that is as a function of $P(T)$ and $P(Data|T)$, it is obvious that the prior and likelihood enable, using 1 to construct a new probability statement about *T* given the data as follows

$$P(Truth|Data) = the\ posterior. \tag{7}$$

The purpose of this chapter was to introduce the concept of Bayesian inference and application to the real world problem such as game theory (Bayesian Game). In this chapter was organized as follows. In Sections 2 and 3, we present Model-based Bayesian inference and the components of Bayesian inference, respectively. The last section contains some applications of Bayesian inference.

2. Model-based Bayesian inference

The basic of Bayesian inference is continued by Bayes' theorem. From (1), replacement *B* with observations *y*, *A* with the set of parameter Θ, and probabilities *P* with densities *p*, results as the following

$$p(\Theta|y) = \frac{p(y|\Theta)p(\Theta)}{p(y)} \tag{8}$$

which $p(y)$ is the marginal likelihood of *y*, $p(\Theta)$ is the set prior distributions of the set of parameter Θ before *y* is observed, $p(y|\Theta)$ is the likelihood of *y* underneath a model and $p(\Theta|y)$ is the joint posterior distribution of Θ that expresses uncertainty about parameter set Θ after taking both the prior and data into system. Because there are often multiple parameters, Θ presents a set of *j* parameters, we have

$$\Theta = \theta_1, \theta_2, ..., \theta_j.$$

The term

$$p(y) = \int p(y|\Theta)p(\Theta)d\Theta \qquad (9)$$

determines the *marginal likelihood* (or the *prior predictive distribution*) of y which it was introduced by Jeffreys [2], and may be set to c where c is an unknown constant. This distribution shows what y should be similar to given the model, before y has been observed. Only the prior probabilities and the model's likelihood function are applied for $p(y)$. The presence of $p(y)$ normalizes the joint posterior distribution, $p(\Theta|y)$ guarantee it is a proper distribution and integrates to 1. From replacement $p(y)$ with a *constant of proportionality c*, the Bayes' theorem becomes to

$$p(\Theta|y) = \frac{p(y|\Theta)p(\Theta)}{c}. \qquad (10)$$

We get

$$p(\Theta|y) \propto p(y|\Theta)p(\Theta) \qquad (11)$$

when $\propto$ is *proportional to*.

This formulation (11) be shown as the unnormalized joint posterior being proportional to the likelihood multiply with the prior. Howsoever, the aim of this model is often not to concluding the non-normalized joint posterior distribution, however to concluding the marginal distributions of the parameters. The set of all Θ can partitioned as

$$\Theta = \{\Phi, \Lambda\} \qquad (12)$$

when the interest sub-vector denote by Φ and the complementary sub-vector of Θ denoted by Λ, usually called to as a vector of nuisance parameters. For a Bayesian scope, the presence of nuisance parameters does not pose any formal, theoretical problems. A nuisance parameter is a parameter that exists in the joint posterior distribution of a model, though it is not a interest parameter. The marginal posterior distribution of ϕ, the interest parameter, can be shown as

$$p(\phi|y) = \int p(\phi, \Lambda|y)d\Lambda. \qquad (13)$$

In model-based Bayesian inference, Bayes' theorem is applied to approximate the non-normalized joint posterior distribution, and lastly the user can evaluate and make inferences by the marginal posterior distributions.

3. The components of Bayesian inference

In this section, we presents about the components of Bayesian inference which contains the prior distributions, the likelihood or likelihood function and the joint posterior distribution as follows.

1. $p(\Theta)$ is the prior distributions for set of Θ, and uses probability as a methods of quantifying uncertainty about Θ before taking the data into system.

2. $p(y|\Theta)$ is the function of likelihood which all variables are associated in a full probability model.

3. $p(\Theta|y)$ is the joint posterior distribution that shows uncertainty about Θ after taking both the prior and the data into system. If Θ is partitioned into a single parameter of interest ϕ and the remaining parameters are considered nuisance parameters, then the marginal posterior distribution of ϕ denote by $p(\phi|y)$.

3.1. Prior distribution

The prior distribution is a main concept of Bayesian and shows the information about an uncertain Θ that is merged with the probability distribution of new data to yield the posterior distribution which in turn is applied for future inferences and decisions about Θ. The existence of a prior distribution for any problem can justified by some axioms of decision theory; which we now focus for how to set up a prior distribution for every given application. Generally, Θ will be a vector, but for easiness we will point as on $p(\Theta)$.

By well-identified and large sample sizes, suitable alternatives of $p(\Theta)$ will have minor effects on posterior inferences. This definition might look like to be circular, but in practice one can check the dependence on $p(\Theta)$ by a sensitivity analysis: comparing posterior inferences under different suitable alternatives of $p(\Theta)$.

If the sample size is small, or available data provide only indirect information about the parameters of interest, then $p(\Theta)$ becomes more important. In various cases, nevertheless, models can be set up hierarchically, such that clusters of parameters have shared $p(\Theta)$, which can themselves be approximated from data. Prior probability distributions have belonged to one of two kinds as informative and uninformative priors. In this section, four kinds of priors which include informative, weakly informative, least informative, and uninformative, are shown according to information and the aim in the use of the prior.

3.1.1. Informative prior

If prior information is obtainable about Θ, it should be included in $p(\Theta)$. If the current model is homologous to a previous model, and the current model is goal to be an adjusted version dependent on more current data, then the posterior distribution of Θ from the previous model maybe used as $p(\Theta)$ for the current model.

Now, every version of a model is not start from scratch, based only on the current data, but the cumulative effects of all data, past and current, can be taken into system. To sure the current data do not dominate the prior, in 2000, Ibrahim and Chen [3] presented the power prior which it is a class of informative prior distribution that takes early data and results into system. If the current data is very homologous to the previous data, then the precision of the posterior distribution increases when including more information from previous models. If the current

data differs tremendously, then the posterior distribution of Θ maybe in the tails of the prior distribution for Θ, therefore $p(\Theta)$ contributes less density in its tails.

Sometimes informative prior is not ready to be applied, for example when it resides in other person, as in an expert. For this way, their human personal beliefs of the probability for the event must be elicited into the form of a suitable probability density function which this process is said to be prior elicitation.

3.1.2. Weakly informative prior

Weakly informative prior (in the short term: WIP) use prior information for regularization and stabilization, providing sufficient prior information to prevent results that contradict our knowledge for example an algorithmic failure to explore the state space. Other aim is for WIPs to use less prior information than is really available. WIPs should provide some of the useful of prior information while avoiding some of the risk from using information which does not exist. WIPs are the most common priors in practice and are liked by subjective Bayesians.

Selecting WIPs may be cumbersome. WIPs distributions should change with the sample size, since a model should have sufficient prior information to learn from the data, but the prior information must also be weak sufficient to learn from the data.

In practice, this is an example of WIPs. It is favor, for well reasons, to center and scale all continuous predictors [4]. Though centering and scaling predictors is not talked about here, but it should be clear that the potential range of the posterior distribution of θ for a centered and scaled predictor should be small. A favor WIPs for a centered and scaled predictor may be $\theta \sim \mathcal{N}(0, 10,000)$ where θ is normal distribution agreeable to a mean of 0 and a variance of 10,000. Here, the density for θ is nearly flat. Nonetheless, the fact that it is not perfectly at yields well properties for numerical estimation algorithms. In both Bayesian and Frequentist inference, it is possible for numerical estimation algorithms to become stuck in regions of at density which become more common as sample size decreases or model complexity increases. Numerical estimation algorithms in Frequentist inference function as though a at prior were used, thus numerical estimation algorithms in Frequentist inference become stuck more frequently than numerical estimation algorithms in Bayesian inference. Prior distributions that are not completely at allow sufficient information for the numerical estimation algorithm to continue to diagnose the goal density, the posterior distribution.

After updating a model in which WIPs exist, the user should be investigating the posterior. If the posterior contradicts knowledge, then the WIPs must be revised by including information that will make the posterior consistent with knowledge [4]. A favor purpose Bayesian criticism against WIPs is that there is no precise mathematical form to derive the optimal WIPs for a given model and data.

3.1.2.1. Vague priors

A vague prior, is said to be a diffuse prior which difficult to define, after considering WIPs. In 2005, Lambert, Sutton, Burton, Abrams and Jones introduce the first formal move from vague

to WIPs. After conjugate priors were introduced by Raiffa and Schlaifer in 1961 which most applied Bayesian has applied vague priors, parameterized to estimate the idea of uninformative priors.

Normally, a vague prior is a conjugate prior together with a large size parameter. Howsoever, if the sample size is small then vague priors may be problems. All most problems about vague priors and small sample size are implicated with scale rather than location. The problem can be particularly acute in random-effects models which it is used rather loosely in this here to imply exchangeable, hierarchical and multilevel structures. A vague prior is defined as commonly being a conjugate prior that is intent to estimate an uninformative prior and without two goals of regularization and stabilization.

3.1.3. Least informative prior

Least informative priors (for short term LIP) is applied here to describe a class of prior in which the aim is to minimize the amount of subjective information content, and to apply a prior that is determined only by the model and observed data. The rationale for using LIPs is often called *to let the data speak for themselves*. LIPs are preferred by objective Bayesians. LIPs are contains Flat Priors [12], Hierarchical Prior [4], Jeffreys Prior [2], MAXENT [5] and Reference Priors [6–8] etc.

3.1.4. Uninformative prior

Traditionally, most of the above descriptions of prior distributions were classified as uninformative priors. However, uninformative priors do not really exist (see in [9]) and all priors are informative in some ways. Moreover, there have been various names associated with uninformative priors including diffuse, minimal, non-informative, objective, reference, uniform, vague, and perhaps weakly informative etc.

3.1.5. Proper and improper priors

It is important for the prior distribution to be proper. A prior distribution, $p(\theta)$, is improper when $\int p(\theta)d\theta = \infty$.

Before, an unbounded uniform prior distribution is an inappropriate prior distribution since $p(\theta) \propto 1$, for $\theta \in (-\infty, \infty)$. An inappropriate prior distribution may be cause an inappropriate posterior distribution. If the posterior distribution is inappropriate, then inferences are invalid.

To determine the propriety of a joint posterior distribution, the marginal likelihood should be finite for any y. Again, the marginal likelihood is $p(y) = \int p(y|\Theta)p(\Theta)d\Theta$. Although inappropriate prior distributions can be applied, it is good practice to avoid them.

3.2. Likelihood

To completely for the definition of a Bayesian, both the prior distributions and the likelihood must be estimated or completely specified. The likelihood or $p(y|\Theta)$, contains the available information provided by the sample. The likelihood is $p(y|\Theta) = \prod_{i=1}^{n} p(y_i|\Theta)$.

The data y effect to the posterior distribution $p(\Theta \mid y)$ only through the likelihood $p(\Theta \mid y)$. In this way, Bayesian inference believes the likelihood principle which states that for a given sample of data, any two probability models $p(\Theta \mid y)$ that have the same likelihood yield the same inference for Θ.

3.3. Posterior distribution

Recent theoretical and applied overviews of Bayesian statistics, including many examples and uses of posterior distributions, see [10–12]. The posterior distributions for decision-making about home radon exposure are discussed in [13].

The posterior distribution summarizes the current state of knowledge about all the uncertain quantities in a Bayesian analysis. Analytically, the posterior density is the product of the prior density and the likelihood. In a complicated analysis, the joint posterior distribution can be summarized by a set of L simulation draws of the vector of uncertain quantities $w_1, w_2, \ldots, w_J$, as illustrated in the following matrix:

l	w_1	w_2	...	w_J
1	.	.	...	.
2	.	...	...	.
...	...	...	...	.
L	.	.	...	.

The marginal posterior distribution for any unknown quantity w_l can be summarized by its column of L simulation draws. In many examples it is not necessary to construct the entire table ahead of time; rather, one creates the L vectors of posterior simulations for the parameters of the model and then uses these to construct posterior simulations for other unknown quantities of interest, as necessary.

4. Application to games theory

In this section, we present the application of Bayesian inference to the real world problems such as Bayesian Game as follows.

4.1. The classical games

The basic contents of the n-person game was presented by John Forbes Nash [14] in 1950. Also, he first shows the existence of equilibrium for this model when the player's preferences are representable by continuous quasi-concave utilities and the sets of strategy are simplex. The definition of an n-person game can be written as below.

Definition 4.1

The normal form of an $n-$ person game is $(X_i, r_i)_{i=1}^n$, where for each $i \in \{1, 2, \ldots, n\}$, the set of individual strategies of player i denoted by X_i which X_i is a non-empty set and r_i is the preference relation on $X := \prod_{i \in I} X_i$ of player i.

The individual preferences r_i are usually represented by utility functions, i.e. for each $i \in \{1, 2, \ldots, n\}$ there exist a real valued function $u_i : X := \prod_{i \in I} X_i \to \mathbb{R}$ such that:

$$x r_i y \Leftrightarrow u_i(x) \geq u_i(y), \forall x, y \in X.$$

Then the normal form of $n-$ person game is transformed to $(X_i, u_i)_{i=1}^n$.

The solution of this game is called Nash equilibrium as below.

Definition 4.2

The Nash equilibrium for the game $(X_i, u_i)_{i=1}^n$ is a point $x^* \in X$ which satisfies for each $i \in \{1, 2, \ldots, n\} : u_i(x^*) \geq u_i(x^*, x_i)$ for each $x_i \in X_i$.

The following theorem offers sufficient conditions for the existence of Nash equilibrium.

Theorem 4.3

Let $\Gamma = (X_i, u_i)_{i=1}^n$ be a n-person game and denoted by f the real-valued function on $X \times X$ defined by $f(x, y) = \Sigma_{i=1}^n u_i(x_{-i}, y_i)$. Let us assume that

1. *for each $i \in \{1, 2, \ldots, n\}$, X_i is a non-empty compact convex subset of a Hausdorff linear topological space;*

2. *for each $i \in \{1, 2, \ldots, n\}$, $u_i(\cdot, x_i)$ is continuous on $X_{-i} = \prod_{i \neq j} X_j$ for each fixed $x_i \in X_i$;*

3. *$\Sigma_{i=1}^n u_i$ is continuous on X;*

4. *$f(x, \cdot)$ is quasi-concave on X, for each $x \in X$.*

Then, Γ has an equilibrium.

Proof. See in [34].

Next, we present some examples of Nash equilibrium for two persons game as follows.

Example 4.4

The battle of the sexes game has two Nash equilibrium (MT, FT), (MS, FS) with $(3, 2)$ and $(2, 3)$, where "Male like playing tennis" denoted by MT, "Male like shopping" denoted by MS, "Female like playing tennis" denoted by FT and "Female like shopping" denoted by FS, see in **Figure 1**.

Example 4.5

The oligopoly behavior game is a unique Nash equilibrium (Aa, Ba) where "A coffee shop use a strategy for don't advertising" denoted by Ad, "A coffee shop use a strategy for advertising"

Figure 1. The battle of the sexes game.

Figure 2. The oligopoly behavior game.

denoted by *Aa*, "A coffee shop use a strategy for do not advertising" denoted by *Bd*, and "A coffee shop use a strategy for advertising" denoted by *Ba*, see in **Figure 2**.

4.2. The Bayesian games

For a long time, we have been supposed that everything in the game was normal knowledge for everyone playing. However, real players may have private information about their own payoffs, their type or preferences, etc. The way to modeling this situation of asymmetrical information is by recurring to the concept was defined by Harsanyi in 1967. The key is to introduce a move by the nature, which changes the uncertainty by converting an asymmetrical information problem into an imperfect information problem. The concept is the nature moves determining players' types, a concept that collects all the private information relevant them (i.e. payoffs, preferences, beliefs of another players, etc.).

Definition 4.6

The normal form of Bayesian games with incomplete information include:

1. the players $i \in \{1, 2, \ldots, I\}$;

2. the set of finite action for each player $a_i \in A_i$;

3. the finite type set for each player $\theta_i \in \Theta_i$;

4. a probability distribution on types $p(\theta)$

5. $u_i : A_1 \times A_2 \times \ldots \times A_I \times \Theta_1 \times \Theta_2 \times \ldots \times \Theta_I \rightarrow \mathbb{R}$, where u_i is utilities function.

It is important to discuss some parts of the definition. Players' types comprise all relevant information about some player's private characteristics. The type of θ_i is only observed by player i who uses this information both to make decisions and to update itself beliefs about the likelihood of opponents' types.

Combining actions and types for each player it is possible to create the strategies. Strategies will be given by $s_i : \Theta_i \rightarrow A_i$, with elements $s_i(\theta_i)$ where Θ_i is the type space and A_i is the action space. A strategy may determine different actions to different types. Lastly, utilities are computed by each player by taking expectations over types using itself own conditional beliefs about opponents' types. Hence, if player i uses the pure strategy s_i, other players use the strategies s_i and player i's type is θ_i, the expected utility can be presented as follows

$$Eu_i(s_i|s_{-i}, \theta_i) = \sum_{\theta_{-i} \in \Theta_{-i}} u_i(s_i, s_{-i}(\theta_{-i}), \theta_i, \theta_{-i}) p(\theta_{-i}|\theta_i).$$

A Bayesian Nash Equilibrium (for short term: BNE) is basically the same concept than a Nash Equilibrium with the addition that players need to take expectations over opponents' types as follows.

Definition 4.7

A Bayesian Nash Equilibrium is a Nash Equilibrium of a Bayesian Game, i.e. $Eu_i(s_i|s_{-1}, \theta_i) \geq Eu_i(s'_i|s_{-i}, \theta_i)$ for all $s'_i \in S_i$ and for all types θ_i occurring with positive probability.

The following theorem for the existence of Bayesian Nash Equilibrium.

Theorem 4.8

Every finite Bayesian Games has a Bayesian Nash Equilibrium.

Example 4.9

Consider the Bayesian games as follows:

1. Nature decides that the payoffs are as in matrix I or II, with probabilities;

2. ROW is informed of the choice of nature but COL is not;

3. The choices of ROW include U or D and the choices of COL include L or R where these choices are made simultaneously;

4. Payoffs are as in the matrix chosen from nature.

For any of the Bayesian games, we will find all BNE. All equilibrium in mixed behavioral strategies can be written as.

Matrix I:

	L	R
U	(1, 1)	(0, 0)
D	(0, 0)	(0, 0)

Matrix II:

	L	R
U	(0, 0)	(0, 0)
D	(0, 0)	(2, 2)

4.2.1. Pure strategy BNE

First, we will deflate the case of incomplete information problem as a static extended game with all of possible strategies: $\widehat{\Gamma}$. It can be presented follow Harsanyi, that the Nash Equilibrium of $\widehat{\Gamma}$ is the same equilibrium of the imperfect game presented. The idea is to deflate a game such that all the ways the game can follow is considered in the extended game $\widehat{\Gamma}$.

The first step is to define the strategies for all player.

Since he does not know in which matrix the game is played, so, COL has only two strategies which contain L and R.

ROW knows in which Matrix the game occurs, and the strategies are UU, UD, DU and DD where UD is played U in Matrix I and D in Matrix II.

The probability knowledge, the nature locates the game in any matrix. The new extended game $\widehat{\Gamma}$ can be shown as:

	L	R
UU	$\left(\frac{1}{2}, \frac{1}{2}\right)$	(0, 0)
UD	$\left(\frac{1}{2}, \frac{1}{2}\right)$	(1, 1)
DU	(0, 0)	(0, 0)
DD	(0, 0)	(1, 1)

Remember that DU is a dominated strategy for ROW. After displacement that possibility, the game has 3 pure Nash Equilibrium as follows $\{(UU; L); (UD; R); (DD; R)\}$.

4.2.2. Mixed strategy BNE

Sequent to obtain the mixed strategies we will make another kind of analysis and try to repeat the three pure BNE obtained before.

Suppose the probabilities of playing each action are as displayed in the matrices as below, where y is the probability COL plays L, if the game is in Matrix I then x is the probability ROW plays U and if the game is in Matrix II then z is the probability ROW plays U.

4.2.3. Player's best respones

- In Matrix I: we get ROW's best response as follows

  ROW would play U, $x=1$, if $1y + 0(1-y) > 0$, then $y > 0$, which can be concluded as:

  a. if $y > 0$, then $x = 1$;

  b. if $y = 0$, then $x \in [0, 1]$.

- In Matrix II: we get ROW's best response as follows

  ROW would play D, $z=0$ if $0 < 2(1-y)$ then $y < 1$ which can be concluded as:

  c. if $y < 1$, then $z = 0$;

  d. if $y = 1$ then $z \in [0, 1]$.

- In Matrix I and II: we get COL's best response as follows

  COL would play L, $y = 1$ if

$$\frac{1}{2}[1x + 0(1-x)] + \frac{1}{2}[0z + 0(1-z)] > \frac{1}{2}[0x + 0(1-x)] + \frac{1}{2}[0z + 2(1-z)]$$

  then $x > 2(1-z)$ which can be summarized as:

  e. if $x = 2(1-z)$, then $y \in [0, 1]$;
  f. if $x > 2(1-z)$, then $y = 1$;
  g. if $x < 2(1-z)$, then $y = 0$.

Next, we can check each the possibilities in order to find the Nash Equilibrium, such as those strategies stable for any players. Let us start by checking COL's strategies since there are less combinations.

4.2.4. Mixed equilibrium

Case 1:

If $y = 0$, we have b. $x \in [0, 1]$ and c. $z = 0$. Here, we want to check this is a equilibrium from COL's point of view. By g., we can see that if $z = 0$, then $x < 2$ which always hold and that $y = 0$.

This Nash equilibrium supports two of the three pure BNE found before: (DD, R), which is the same as $y = 0$, $x = 0$ and $z = 0$ and (UD, R) which is the same as $y = 0$, $x = 1$ and $z = 0$.

Thus, we get Nash equilibrium of $y = 0$, $x \in [0, 1]$ and $z = 0$.

There are many BNE in which column plays R and row plays $xU + (1-x)D$, when $x \in [0, 1]$ if Matrix I occurs and D if Matrix II occurs.

Case 2:

If $y = 0$, we have *d.* $z \in [0, 1]$ and from *a.* $x = 1$.

From *f.*, we can see that when $x = 1$, then it should be the case that $z \geq \frac{1}{2}$ in order to be true that $y = 1$. Hence, these BNE are restricted to $y = 1$, $z \in \left[\frac{1}{2}, 1\right]$ and $x = 1$.

This BNE supports the third pure Nash Equilibrium found before: (*UU, L*), which is the same as $y = 1$, $x = 1$ and $z = 1$.

There are many BNE in which column plays *L* and row plays *U* if Matrix I occurs and $zU + (1 - z)D$, where $z \in \left[\frac{1}{2}, 1\right]$ if Matrix II occurs.

Case 3:

If $y \in (0, 1)$, we have *a.* $x = 1$ and *c.* $z = 0$. By *e.*, we can see that in order y belongs to $[0, 1]$ it should be the case that $x = (1 - z)$. However it is impossible this equality to hold if both $z = 0$ and $x = 1$.

Therefore, the case if $y \in (0, 1)$ is not a Bayesian Nash equilibrium.

4.3. Abstract economy model

Later, the existence of social equilibrium was proved Debreu [15]. Also Arrow and Debreu [16] proved the existence of Walrasian equilibrium. The classical abstract economy game introduced by Shafer and Sonnenschein [17] or Borglin and Keiding [18] consists of a finite set of agents, each characterized by certain constraints and preferences, explained by correspondences. Following many previous authors ideas, they studied the existence of equilibrium for generalized games (see, for example, [19–27] and the references therein). Now, we show some definitions of an abstract economy model and equilibrium of this model as follows. Let the set of agents be the finite set $\{1, 2, \ldots, n\}$. For each $i \in \{1, 2, \ldots, n\}$ let X_i be a non-empty set.

Definition 4.10

An abstract economy $\Gamma = (X_i, A_i, P_i)_{i=1}^{n}$ is defined as a family of n ordered triplets (X_i, A_i, P_i), where for each $i \in I$:

1. $A_i : \prod_{i \in I} X_i \to 2^{X_i}$ is constraint correspondence and

2. $P_i : \prod_{i \in I} X_i \to 2^{X_i}$ is preference correspondence.

Definition 4.11

An equilibrium for Γ is a point $x^* \in \prod_{i \in I} X_i$ which satisfies for each $i \in \{1, 2, \ldots, n\}$:

1. $x^* \in A_i(x^*)$;

2. $A_i(x^*) \cap P_i(x^*) = \varnothing$.

Theorem 4.12

Let $\Gamma = (X_i, A_i, P_i)_{i=1}^{n}$ be an abstract economy which satisfies, for each $i \in \{1, 2, \ldots, n\}$:

1. X_i is a non-empty compact convex subset in $\mathbb{R}^l$;

2. A_i is a continuous correspondence;

3. for each $x \in X$, $A_i(x)$ is non-empty compact and convex;

4. P_i has an open graph in $X \times X_i$ and for each $x \in X$, $P_i(x)$ is convex;

5. for each $x \in X$, $x_i \notin P_i(x)$.

Then, Γ has an equilibrium.

Proof. See in [34].

4.4. Fuzzy games

The first concept of a fuzzy set was introduced by Zadeh [28] in 1965. Fuzzy set theory has been shown to be a gainful tool to describe situations in which the data are imprecise or vague. The theory of fuzzy sets has become a well framework for studying results concerning fuzzy equilibrium existence for abstract fuzzy economies. The first study of a fuzzy abstract economy (or a fuzzy game) has been studied by Kim and Lee in [29], they shown the existence of the equilibrium for 1-person fuzzy game. Also Kim and Lee [29] shown the existence of equilibrium for generalized games when the constraints or preferences are vague due to the agent's behavior. In 2009, Patriche [30] studied the Bayesian abstract economy game and proved the existence of equilibrium for an abstract economy game with differential information and a measure space of agents. However, the existence of random fuzzy equilibrium for fuzzy game has not been studied so far. In 2013, Patriche [31] defined the Bayesian abstract economy game and proved the existence of the Bayesian fuzzy equilibrium for this game. Also, Patriche [32] defined the new Bayesian abstract fuzzy economy game and proved the existence of the Bayesian fuzzy equilibrium for this game which it is characterized by a private information set, an action fuzzy mapping, a random fuzzy constraint one and a random fuzzy preference mapping. Recently, Patriche [33] defined the fuzzy games and applications to systems of generalized quasi-variational inequalities problem. The Bayesian fuzzy equilibrium concept is an extension of the deterministic equilibrium. She also generalized and extended the former deterministic models introduced by Debreu [15], Shafer and Sonnenschein [17] and Patriche [34]. Very recently, Saipara and Kumam [35] introduced the model of general Bayesian abstract fuzzy economy for product measurable spaces, and proved the existence for Bayesian fuzzy equilibrium of this model as follows.

For each $i \in I$, let $(\Omega_i, \mathcal{Z}_i)$ be a measurable space, $(\Omega, \mathcal{Z})$ be the product measurable space where $\Omega := \prod_{i \in I} \Omega_i$, $\mathcal{Z} := \otimes_{i \in I} \mathcal{Z}_i$ and μ is a probability measure on $(\Omega, \mathcal{Z})$. Let Y denote the strategy or commodity space, where Y is a separable Banach space.

Let I be a non-empty finite set (the set of agents). For each $i \in I$, let $X_i : \Omega_i \to \mathcal{F}(Y)$ be a fuzzy mapping, and let $z_i \in (0, 1]$.

Let $L_{X_i} = \{x_i \in S(X_i(\cdot))_{z_i} : x_i$ is Σ_i-measurable$\}$. Denote by $L_X = \prod_{i \in I} L_{X_i}$ and by the set $\prod_{i \neq j} L_{X_j}$. An element x_i of L_{X_i} is called a strategy for agent i. The typical element of L_{X_i} is denoted by x_i and

that of $(X_i(\omega_i))_{z_i}$ by $x_i(\omega_i)$ (or x_i). We can define a general Bayesian abstract fuzzy economy model of product measurable spaces as follow.

Definition 4.13

A general Bayesian abstract fuzzy economy model of product measurable spaces is defined as follows:

$$\Gamma = \left(\left((\Omega_i, \mathcal{Z}_i)_{i \in I}, \mu \right), \left(X_i, \Sigma_i, (A_i, a_i), (P_i, p_i), z_i \right)_{i \in I} \right),$$

where I is non-empty finite set (the set of agents) and:

a. $X_i : \Omega_i \to \mathcal{F}(Y)$ *is a action (strategy) fuzzy mapping of agent i;*

b. Σ_i *is a sub σ-algebra of $\mathcal{Z} = \otimes_{i \in I} \mathcal{Z}_i$, which denotes the private information of agent i;*

c.   *for each $\omega_i \in \Omega_i$, $A_i(\omega_i, \cdot) : L_X \to \mathcal{F}(Y)$ is the random fuzzy constraint mapping of agent i;*

d.   *for each $\omega_i \in \Omega_i$, $P_i(\omega_i, \cdot) : L_X \to \mathcal{F}(Y)$ is the random fuzzy preference mapping of agent i;*

e. $a_i : L_X \to (0, 1]$ *is a random fuzzy constraint function, and $p_i : L_X \to (0, 1]$ is a random fuzzy preference function of agent i;*

f. $z_i \in (0, 1]$ *is such that for all $(\omega_i, x) \in \Omega_i \times L_X$, $(A_i(\omega_i, \tilde{x}))_{a_i(\tilde{x})} \subset (X_i(\omega_i))_{z_i}$ and $(P_i(\omega_i, \tilde{x}))_{p_i(\tilde{x})} \subset (X_i(\omega_i))_{z_i}$.*

The Bayesian fuzzy equilibrium for a general Bayesian abstract fuzzy economy model of product measurable spaces is defined as follows.

Definition 4.14

A Bayesian fuzzy equilibrium for Γ is a strategy profile $\tilde{x}^* \in L_X$ such that for all $i \in I$,

i. $\tilde{x}^*(\omega_i) \in cl(A_i(\omega_i, \tilde{x}^*))_{a_i(\tilde{x}^*)}$ $\mu - a.e.;$

ii. $(A_i(\omega_i, \tilde{x}^*))_{a_i(\tilde{x}^*)} \cap (P_i(\omega_i, \tilde{x}^*))_{p_i(\tilde{x}^*)} = \varnothing$ $\mu - a.e..$

Theorem 4.15

Let I be a non-empty finite set. Let the family

$\Gamma = \left(\left((\Omega_i, \mathcal{Z}_i)_{i \in I}, \mu \right), \left(X_i, \Sigma_i, (A_i, a_i), (P_i, p_i), z_i \right)_{i \in I} \right)$ *be a general Bayesian abstract economy model of product spaces satisfy (a)-(j). Then, there exists a Bayesian fuzzy equilibrium for Γ.*

For each $i \in I$, the following conditions are sastisfied:

a. $X_i : \Omega_i \to \mathcal{F}(Y)$ *is such that $\omega_i \to X_i(\omega_i)_{z_i} : \Omega_i \to 2^Y$ is a non-empty convex weakly compact-valued and integrably bounded correspondence;*

b. $X_i : \Omega_i \to \mathcal{F}(Y)$ *is such that $\omega_i \to X_i(\omega_i)_{z_i} : \Omega_i \to 2^Y$ is $\sum_i -$ lower measurable;*

c.   *For each $(\omega_i, \tilde{x}) \in \Omega_i \times L_X$, $(A_i(\omega_i, \tilde{x}))_{a_i(\tilde{x})}$ is convex and has a non-empty interior in the relative norm topology of $(X_i(\omega_i))_{z_i}$;*

d. *the correspondence* $(\omega_i, \tilde{x}) \to (A_i(\omega_i, \tilde{x}))_{a_i(\tilde{x}):\Omega_i \times L_X \to 2^Y}$ *has a measurable graph, i.e.,*

$\left\{ (\omega_i, \tilde{x}, y) \in \Omega_i \times L_X \times Y : y \in (A_i(\omega_i, \tilde{x}))_{a_i(\tilde{x})} \right\} \in \mathcal{F}_i \otimes \mathfrak{B}(L_X) \otimes \mathfrak{B}(Y),$ *where* $\mathfrak{B}_{\omega_i}(L_X)$ *is the Borel* $\sigma-$ *algebra for the weak topology on* L_X *and* $\mathfrak{B}(Y)$ *is the Borel* $\sigma-$ *algebra for the norm topology on* Y;

e. *the correspondence* $(\omega_i, \tilde{x}) \to (A_i(\omega_i, \tilde{x}))_{a_i(\tilde{x})}$ *has weakly open lower sections, i.e., for each* $\omega_i \in \Omega_i$

and for each $y \in Y$, *the set* $\left((A_i(\omega_i, \tilde{x}))_{a_i(\tilde{x})} \right)^{-1} (\omega_i, \tilde{x}) = \left\{ \tilde{x} \in L_X : y \in (A_i(\omega_i, \tilde{x}))_{a_i(\tilde{x})} \right\}$ *is weakly open in* L_X;

f. *For each* $\omega_i \in \Omega_i, \tilde{x} \to cl(A_i(\omega_i, \tilde{x}))_{a_i(\tilde{x})} : L_X \to 2^Y$ *is upper semicontinuous in the sense that the set* $\left\{ \tilde{x} \in L_X : cl(A_i(\omega_i, \tilde{x}))_{a_i(\tilde{x})} \right\}$ *is weakly open in* L_X *for every norm open subset* V *of* Y;

g. *the correspondence* $(\omega_i, \tilde{x}) \to (P_i(\omega_i, \tilde{x}))_{p_i(\tilde{x})} : \Omega_i \times L_X \to 2^Y$ *has open convex values such that* $(P_i(\omega_i, \tilde{x}))_{p_i(\tilde{x})} \subset (X(\omega_i))_{z_i}$ *for each* $(\omega_i, \tilde{x}) \in \Omega_i \times L_X$;

h. *the correspondence* $(\omega_i, \tilde{x}) \to (P_i(\omega_i, \tilde{x}))_{p_i(\tilde{x})} : \Omega_i \times L_X \to 2^Y$ *has a measurable graph;*

i. *the correspondence* $(\omega_i, \tilde{x}) \to (P_i(\omega_i, \tilde{x}))_{p_i(\tilde{x})} : \Omega_i \times L_X \to 2^Y$ *has weakly open lower sections, i.e. for each* $\omega_i \in \Omega_i$ *and for each* $y \in Y$, *the set* $((P_i(\omega_i, \tilde{x}))_{p_i(\tilde{x})})^{-1}(\omega_i, y) = \{ \tilde{x} \in L_X : y \in (P_i(\omega_i, \tilde{x}))_{p_i(\tilde{x})} \}$ *is weakly open in* L_X;

j. *For each* $\tilde{x}_i \in L_{X_i}$, *for each* $\omega_i \in \Omega_i, \tilde{x}_i \notin (A_i(\omega_i, \tilde{x}))_{a_i(\tilde{x})} \cap (P_i(\omega_i, \tilde{x}))_{p_i(\tilde{x})}.$

Proof. See in [35].

Moreover, in 1960, Fichera and Stampacchia first introduced the variational inequalities problem, this issue has been widely studied. Next, the basic concept of variational inequalities for fuzzy mappings was first introduced by Chang and Zhu [36] in 1989. In the topic of variational inequalities problem, there are many mathematicians who studied this topic (see, for example, [37, 38]). In 1993, the concept of a random variational inequality was introduced by Noor and Elsanousi [39]. Recently, Patriche [31] used the model of the Bayesian abstract fuzzy economy to prove the existence of solution for the two types of random quasi-variational inequalities with random fuzzy mappings.

5. Conclusion

The main objectives of this chapter was introduced the concept of Bayesian inference and application to some real world problems. In this chapter, we were presented about the basic concept of Bayesian inference which it can be application to the Bayesian game and a general Bayesian abstract fuzzy economy game or a fuzzy game. For application to Bayesian game, we

were shown the solution of Bayesian Nash Equilibrium (BNE) for a Bayesian game with examples. Finally, we were shown the existence of Bayesian fuzzy equilibrium for a fuzzy game.

Acknowledgements

Author details

Wiyada Kumam[1], Plern Saipara[2] and Poom Kumam[3*]

*Address all correspondence to: poom.kum@kmutt.ac.th

1 Rajamangala University of Technology Thanyaburi (RMUTT), Thailand

2 Rajamangala University of Technology Lanna Nan (RMUTL), Thailand

3 King Mongkut's University of Technology Thonburi (KMUTT), Thailand

References

[1] Stigler S. Who discovered Bayes's theorem. The American Statistician. 1983;**37**(4):290-296

[2] Jeffreys H. Theory of Probability. 3rd ed. Oxford, England: Oxford University Press; 1961

[3] Ibrahim J, Chen M. Power prior distributions for regression models. Statistical Science. 2000;**15**:46-60

[4] Gelman A. Scaling regression inputs by dividing by two standard deviations. Statistics in Medicine. 2008;**27**:2865-2873

[5] Jaynes E. Prior probabilities. IEEE Transactions on Systems Science and Cybernetics. 1968;**4**(3):227-241

[6] Berger J, Bernardo J, Dongchu S. The formal definition of reference priors. Annals of Statistics. 2009;**37**(2):905-938

[7] Bernardo J. Reference posterior distributions for Bayesian inference (with discussion). Journal of the Royal Statistical Society, B. 1979;**41**:113-147

[8] Bernardo J. Reference analysis. In: Dey D, Rao C, editors. Handbook of Statistics. Vol. 25. Amsterdam: Elsevier; 2005. p. 17-90

[9]   Irony T, Singpurwalla N. Noninformative priors do not exist: A discussion with Jose M. Bernardo. Journal of Statistical Inference and Planning. 1997;**65**:159-189

[10] Bernardo JM, Smith AFM. Bayesian Theory. New York: Wiley; 1994

[11] Carlin BP, Louis TA. Bayes and Empirical Bayes Methods for Data Analysis. London: Chapman and Hall; 1996

[12] Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian Data Analysis. London: Chapman and Hall; 1995

[13] Lin CY, Gelman A, Price PN, Krantz DH. Analysis of local decisions using hierarchical modeling, applied to home radon measurement and remediation (with discussion). Statistical Science. 1999;**14**:305-337

[14] Nash J. Equilibrium points in *n*-person games. Proceedings of the National Academy of Sciences of the United States of America. 1950;**36**(1):48-49

[15] Debreu G. A social equilibrium existence theorem. Proceedings of the National Academy of Sciences of the United States of America. 1952;**38**:886-903

[16] Arrow KJ, Debreu G. Existence of an equilibrium for a competitive economy. Econometrica. 1952;**22**:265-290

[17] Shafer W, Sonnenschein H. Equilibrium in abstract economies without ordered preferences. Journal of Mathematical Economics. 1975;**2**:345-348

[18] Borglin A, Keiding H. Existence of equilibrium actions and of equilibrium: A note on the "new" existence theorem. Journal of Mathematical Economics. 1976;**3**:313-316

[19] Huang NJ. Some new equilibrium theorems for abstract economies. Applied Mathematics Letters. 1998;**11**(1):41-45

[20] Kim WK, Tan KK. New existence theorems of equilibria and applications. Nonlinear Analysis: Theory, Methods & Applications. 2001;**47**:531-542

[21] Lin LJ, Chen LF, Ansari QH. Generalized abstract economy and systems of generalized vector quasi-equilibrium problems. Journal of Computational and Applied Mathematics. 2007;**208**:341-353

[22] Briec W, Horvath C. Nash points, Ky Fan inequality and equilibria of abstract economies in Max-Plus and B-convexity. Journal of Mathematical Analysis and Applications. 2008;**341**:188-199

[23] Ding XP, Wang L. Fixed points, minimax inequalities and equilibria of noncompact abstract economies in FC-spaces. Nonlinear Analysis: Theory, Methods & Applications. 2008;**69**:730-746

[24] Kim WK, Kum SH, Lee KH. On general best proximity pairs and equilibrium pairs in free abstract economies. Nonlinear Analysis: Theory, Methods & Applications. 2008;**68**:2216-2227

[25] Lin LJ, Liu YH. The study of abstract economies with two constraint correspondences. Journal of Optimization Theory and Applications. 2008;**137**:41-52

[26] Ding XP, Feng HL. Fixed point theorems and existence of equilibrium points of noncompact abstract economies for L_F^*-majorized mappings in FC-spaces. Nonlinear Analysis: Theory, Methods & Applications. 2010;**72**:65-76

[27] Wang L, Cho YJ, Huang NJ. The robustness of generalized abstract fuzzy economies in generalized convex spaces. Fuzzy Sets and Systems. 2011;**176**:56-63

[28] Zadeh LA. Fuzzy sets. Information and Control. 1965;**8**:338-353

[29] Kim WK, Lee KH. Fuzzy fixed point and existence of equilibria of fuzzy games. Journal of Fuzzy Mathematics. 1998;**6**:193-202

[30] Patriche M. Bayesian abstract economy with a measure space of agents. Abstract and Applied Analysis. 2009;**2009**:1-11

[31] Patriche M. Equilibriumof Bayesian fuzzy economies and quasi-variational inequalities with random fuzzy mappings. Journal of Inequalities and Applications. 2013; 374. Article ID 58E35

[32] Patriche M. Existence of equilibrium for an abstract economy with private information and a countable space of actions. Mathematical Reports. 2013;**15**(65)(3):233-242

[33] Patriche M. Fuzzy games with a countable space of actions and applications to systems of generalized quasi-variational inequalities. Fixed Point Theory and Applications. 2014;**2014**(124)

[34] Patriche M. Equilibrium in Games and Competitive Economies. Bucharest: The Publishing House of the Romanian Academy; 2011

[35] Saipara P, Kumam P. Fuzzy games for a general Bayesian abstract fuzzy economy model of product measurable spaces. Mathematical Methods in the Applied Sciences. 2015;**39**(16): 4810-4819

[36] Chang SS, Zhu YG. On variational inequalities for fuzzy mappings. Fuzzy Sets and Systems. 1989;**32**:359-367

[37] Noor MA. Variational inequalities for fuzzy mappings III. Fuzzy Sets and Systems. 2000;**110**:101-108

[38] Park JY, Lee SY, Jeong JU. Completely generalized strongly quasivariational inequalities for fuzzy mapping. Fuzzy Sets and Systems. 2000;**110**:91-99

[39] Noor MA, Elsanousi SA. Iterative algorithms for random variational inequalities. Panamerican Mathematical Journal. 1993;**3**:39-50