Universal Point Estimation, with Applications in Economics, Business and Decision Sciences

Buu-Chau Truong[1], Thi Diem-Chinh Ho[2,3], Thu-Quang Luu[4] and Michael McAleer[5-9,*]

[1]*Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, Vietnam*

[2]*Faculty of Mathematics and Statistics, University of Sciences, Ho Chi Minh City, Vietnam*

[3]*General Faculty, Binh Duong Economics & Technology University, Binh Duong, Vietnam*

[4]*Faculty of Finance, Banking University of Ho Chi Minh City, Vietnam*

[5]*Department of Finance, Asia University, Taiwan*

[6]*Discipline of Business Analytics, University of Sydney Business School, Australia*

[7]*Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands*

[8]*Department of Economic Analysis and ICAE, Complutense University of Madrid, Spain*

[9]*Institute of Advanced Sciences, Yokohama National University, Japan*

Abstract: Estimation is used widely in numerous disciplines, including Mathematics, Statistics, Economics, Business, and Decision Sciences, among others. Estimation is a process for determining an approximation, which is a value that can be used for a number of purposes, even if input data are sufficient, incomplete, missing or unsecure. In practice, estimation relates to "using the value of a statistic inferred from a sample to estimate the value of a corresponding population parameter". Estimation is usually separated into two categories, namely point estimation and interval estimation. The main purpose of this paper is to present a universal approach to the theory and practice of three methods in statistical inference to obtain point estimates, namely the moment, maximum likelihood, and Bayesian methods. The paper also discusses the advantages and disadvantages of the three universal approaches in practical applications in Economics, Business and Decision Sciences.

Keywords: Universal approach, Maximum likelihood, Moment method, Bayesian method.

1. INTRODUCTION

Estimation is a process for detecting an approximation, which is a value that can be used for a number of purposes, even if input data are insufficient, incomplete, missing, or unsecure. Usually, estimation relates to "using the value of a statistic inferred from a sample to estimate the value of a corresponding population parameter".

In mathematics, an estimate is called an approximation, and is written in the form of larger or smaller quantities of each other. It is not always straightforward to determine the exact values of functions, whether known or unknown. The approximation theorem can help to find simpler functions that are very close to the functions to be determined, and to provide more useful calculation tools.

In statistics, an estimate is a value calculated from a sample and is expected to be a representative value to be determined for the population. A statistically valid and optimal estimator would ideally be unbiased, consistent, (asymptotically) efficient, and robust to changes in the underlying assumptions.

In terms of applications, Hawkins (1976) introduces point estimation of the parameters of piecewise regression models. Bafumi *et al*. (2005) provide practical issues in implementing and understanding Bayesian point estimation. Lehmann and Casella (2006) present a theory of point estimation. Schennach (2007) introduces point estimation with an exponentially tilted empirical likelihood. Harchaoui *et al*. (2010) provide multiple change point estimation with a total variation penalty. Hodges and Lehmann (2012) present some problems in minimax point estimation.

Estimation is typically separated into two categories, namely point estimation and interval estimation. Point estimation is used more frequently than interval estimation because point estimation allows reasonably precise estimates of the parameters or moments that are of interest. Issues related to point estimation are still topical and interesting (see Amiri and Allahyari (2012), Chung *et al*. (2015), Yoo *et al*. (2017) and Goplerud, M. (2019), among others).

*Address correspondence to this author at the Department of Finance, Asia University, Taiwan; Tel: 2332-3456#1837; E-mail: michael.mcaleer@gmail.com

JEL: A10, G00, G31, O32.

Estimation is applied in many fields in practice, and is important in Economics, Finance and Business, among other disciplines, because it helps to determine how large-scale activities may develop, plan distributions for workers, buy materials more efficaciously, estimate project revenues, costs, and profits, and so on.

This paper presents a detailed and comprehensive analysis of the theory, examples and practical applications of point estimation in Economics, Business and Decision Sciences. In particular, the paper provides an overview of universal methods to determine point estimates in statistics.

The remainder of the paper is organized as follows. In Section 3, we present some notation, definitions, and background theory to be used. In Section 3, we present three universal methods to obtain point estimates in statistics. We provide some practical examples to find point estimates in statistics in Section 4. Section 5 discusses the advantages and disadvantages of the three approaches. Practical applications of the universal methods to Economics, Business and Decision Sciences are discussed in Section 6. Some concluding remarks and inferences are presented in the final section.

2. NOTATION, DEFINITIONS, AND FOUNDATIONS

In this section, we present some notation, definitions, and background theory to be used in the following sections. We first introduce the concept of a population.

2.1. Population

A population is a set of analogous items that is of interest. A population can be a group of existing objects or an hypothetical and possibly infinite objects perceived as a generalization from experience. A widespread aim of statistical analysis is to draw inferences related to the population. The elements in the population can typically be described by a random variable *X*. If we obtain the probability density function (PDF) $f(x;\theta)$ of the random variable *X* and the parameter θ that is of interest, then we know its population.

Populations that can be of finite size and defined by any number of characteristics is used in statistics to describe all the items of interest. Each item is the subject of a statistical observation that is defined specifically rather than vaguely. Populations allow inferences to be drawn and conclusions presented and analyzed regarding the characteristics of parameters and moments that are of interest.

2.2. Sample

A random statistical sample taken from the population is a subset of the population that is intended to be analyzed. The ratio of the size of the sample to the size of the population is the sampling fraction. It is analyzed to estimate the population parameters and moments using appropriate sample statistics to draw inferences and conclusions. A set of random variables $X_1, X_2, ..., X_n$ represents the sample with probability density function (pdf), $f(x;\theta)$, that is the same as the population. The observations, or the sample values, of the random variables $X_1, X_2, ..., X_n$ are taken to be $x_1, x_2, ..., x_n$.

2.3. Sample Space

Let *X* be a random variable representing a population with pdf $f(x;\theta)$, where θ is a parameter of interest. The set of all possible values of *X* is defined as the sample space denoted by Ω such that $\Omega = \{X \in R^n : f(x;\theta)\}$, where *n* is the sample size.

2.4. Estimation

Let $X_1, X_2, ..., X_n$ be a random sample from the population such that $X \sim f(x;\theta)$ with θ to be the parameter of interest. All statistics used to predict parameter θ are called estimation functions of θ. The sample statistic, denoted as $\hat{\theta}$, is defined as the estimator of θ, and its empirical value is called the estimate of θ.

All statistical methods use random samples to obtain information regarding the population. The population is typically described by the pdf $f(x;\theta)$. A statistical inference is drawn regarding the population distribution of $f(x;\theta)$ based on the information contained in the sample. A statistical inference is a statement based on a sample regarding information in the population. There are three types of statistical inferences that consist of: estimation, hypothesis testing, both of which are in-sample, and out-of-sample forecasting.

In estimation, one needs to obtain a sample value of the parameter θ of the population distribution, $f(x;\theta)$, from the sample data. One of the key issues is how to obtain the estimation function of the parameter θ of the population. There are three widely-used and universal methods of determining point estimation, namely the moment, maximum likelihood, and Bayesian methods. We will discuss each method in the following section.

3. METHODS TO DETERMINE POINT ESTIMATES

In this section, we present the three universal methods for determining point estimates, namely the moment, maximum likelihood, and Bayesian methods. We first discuss the moment method.

3.1. Moment Method

This method was first introduced by the famous British statistician Karl Pearson in 1902. It remains a classical parameter estimation method that is simple, yet yields consistent (though possibly biased) estimators under minimal assumptions. The estimates obtained by using the moment method are not necessarily sufficient statistics, in that they occasionally fail to take account of all the relevant information from the sample.

Let $X_1, X_2, ..., X_m$ be a random sample from the population *X* with pdf given by $f\left(x;\theta_1,\theta_2,...,\theta_m\right)$ where $\theta_1,\theta_2,...,\theta_m$ are the parameters to be estimated. Let :

$$M_k = E\left(X^k\right) = \int_{-\infty}^{+\infty} x^k f\left(x;\theta_1,\theta_2,...,\theta_m\right)dx, \quad (1)$$

be the *k* 'th population moment, and $M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ be the corresponding *k*'th sample moment, for $k = 1,...,m.$

In order to apply the moment method, one first estimates $\theta_1,\theta_2,...,\theta_m$ by setting the first *m* population moments equal to the first *m* sample moments. The procedure of the moment method is expressed in the following steps:

Step 1: Calculate the first *m* population moments and the first *m* sample moments;

Step 2: Set up the *m* systems of equations by setting the first *m* population moments equal to the first *m* sample moments; and

Step 3: Solve the system of equations to obtain the estimation function.

3.2. Maximum Likelihood Method

In statistics, from the given observations, maximum likelihood estimation (MLE) is the most commonly-used method to estimate the parameters for the statistical model from the given observations for which the likelihood function is maximized. More precisely, let $X_1, X_2, ..., X_n$ be a random sample from population *X* with pdf $f\left(x;\theta\right)$, in which θ are the parameters to be estimated. The likelihood function, $L\left(\theta\right)$, is the joint pdf of the sample, $X_1, X_2, ..., X_n$, and is defined as:

$$L\left(\theta\right) = \prod_{i=1}^{n} f\left(x_i;\theta\right). \quad (2)$$

The value θ that leads to the maximum value of the likelihood function, $L\left(\theta\right)$, is the MLE of θ, and is defined by $\hat{\theta}$ such that:

$$\hat{\theta} = \mathrm{Arg}\sup_{\theta\in\Omega} L\left(\theta\right). \quad (3)$$

In general, the estimates obtained using MLE has the following properties:

Property 1 (Unbiasedness and Efficiency)

Let $\left\{Y_i\right\}$ be a sequence of i.i.d. observations and θ be the parameters to be estimated. If an estimator $\hat{\theta}\left(y\right)$ is unbiased and efficient, then $\hat{\theta}\left(y\right)$ is also the minimum variance unbiased estimator (MVUE).

Property 2 (Consistency)

Let $\left\{Y_i\right\}$ be a sequence of i.i.d. observations and θ be the parameters to be estimated. Then, the MLE of θ is consistent.

Property 3 (Asymptotic Normality)

Let $\left\{Y_i\right\}$ be a sequence of i.i.d. observations and θ be the parameters to be estimated, $Y_i \sim f_\theta\left(y\right)$ and $\hat{\theta}$ is the MLE of θ. It follows that:

$$\sqrt{n}\left(\hat{\theta}-\theta\right) \xrightarrow{d} N\left(0, \frac{1}{-E\left(\frac{\partial^2}{\partial\theta^2}\log f_\theta\left(y\right)\right)}\right).$$

Property 4 (functional transformation)

If $\hat{\theta}$ is the MLE of θ and $g\left(\theta\right)$ is a known function of θ, then $g\left(\hat{\theta}\right)$ is the MLE of $g\left(\theta\right)$.

See Lehmann (2004) and Lehmann and Casella (2006) for the proofs of Properties 1 - 4.

The maximum likelihood estimation procedure can be expressed in the following:

Step 1: Let the random sample $x_1, x_2, ..., x_n$ be obtained from the population *X* with pdf $f(x;\theta)$;

Step 2: Determine the likelihood function of the sample $x_1, x_2, ..., x_n$, as shown in (2);

Step 3: Detect $\hat{\theta}$ such that $L(\hat{\theta})$ reaches a maximum, as shown in (3).

3.3. Bayesian Method

The Bayesian estimator is an estimator that minimizes the posterior expected value of a loss function (that is, the posterior expected loss), such that it maximizes the posterior expectation of a utility function, or the posterior likelihood function.

The Bayesian estimation technique treats the parameter as a random variable. Thus, one can assign a probability distribution to indicate the trust of the actual value of the parameter. This is a subjective or prior distribution that is based on the opinions of experienced researchers, and is established before the data are used. Bayes' Theorem gives the conditional or posterior distribution of the parameters given the data by combining the existing prior information together with the data information.

Theoretically, we let $X_1, X_2, ..., X_n$ be a random sample with pdf $f(x,\theta)$ or $f(x|\theta)$, where θ are the parameters to be estimated. The pdf of θ is defined as the prior distribution of θ and is denoted by $h(\theta)$. The conditional density or posterior distribution, $k(\theta | x_1, x_2, ..., x_n)$, of θ for the sample $x_1, x_2, ..., x_n$ is given by:

$$k(\theta | x_1, x_2, ..., x_n) = \frac{h(\theta)\prod_{i=1}^{n} f(x_i | \theta)}{\int_{-\infty}^{+\infty} h(\theta)\prod_{i=1}^{n} f(x_i | \theta) d\theta}. \quad (4)$$

Let $\hat{\theta}$ be an estimator of θ and $L(\hat{\theta},\theta)$ be the loss function. Then the posterior expected loss is given as:

$$g(\hat{\theta}) = \int_{\Omega} L(\hat{\theta},\theta) k(\theta | x_1, x_2, ..., x_n) d\theta. \quad (5)$$

The Bayesian estimator, $\hat{\theta}$, is the estimator that minimizes the posterior expected loss. In this paper, we consider two types of loss functions, namely quadratic loss $L(\hat{\theta},\theta) = (\hat{\theta} - \theta)$ and absolute loss $L(\hat{\theta},\theta) = |\hat{\theta} - \theta|$. For quadratic loss, the posterior distribution is the mean square error, such that:

$$g(\hat{\theta}) = \int_{-\infty}^{+\infty} (\hat{\theta} - \theta)^2 k(\theta | x_1, x_2, ..., x_n) d\theta. \quad (6)$$

Let $g'(\hat{\theta})$, so that :

$$\hat{\theta}\int_{-\infty}^{+\infty} k(\theta | x_1, x_2, ..., x_n) d\theta = \int_{-\infty}^{+\infty} \theta k(\theta | x_1, x_2, ..., x_n) d\theta$$

such that:

$$\hat{\theta} = \int_{-\infty}^{+\infty} \theta k(\theta | x_1, x_2, ..., x_n) d\theta \quad (7)$$

as

$$\int_{-\infty}^{+\infty} k(\theta | x_1, x_2, ..., x_n) d\theta = 1.$$

For the absolute loss, the posterior distribution is the mean absolute error, as shown in the following:

$$\begin{aligned} g(\hat{\theta}) &= \int_{-\infty}^{+\infty} |\hat{\theta} - \theta| k(\theta | x_1, x_2, ..., x_n) d\theta \\ &= \hat{\theta}\int_{-\infty}^{\hat{\theta}} k(\theta | x_1, x_2, ..., x_n) d\theta - \hat{\theta}\int_{\hat{\theta}}^{+\infty} k(\theta | x_1, x_2, ..., x_n) d\theta \quad (8) \\ &- \int_{-\infty}^{\hat{\theta}} \theta k(\theta | x_1, x_2, ..., x_n) d\theta + \int_{\hat{\theta}}^{+\infty} \theta k(\theta | x_1, x_2, ..., x_n) d\theta. \end{aligned}$$

Set:

$g'(\hat{\theta}) = 0$, we get $\int_{-\infty}^{\hat{\theta}} k(\theta | x_1, x_2, ..., x_n) d\theta = \int_{\hat{\theta}}^{+\infty} k(\theta | x_1, x_2, ..., x_n) d\theta.$

In addition, we have:

$$\int_{-\infty}^{\hat{\theta}} k(\theta | x_1, x_2, ..., x_n) d\theta + \int_{\hat{\theta}}^{+\infty} k(\theta | x_1, x_2, ..., x_n) d\theta = 1.$$

Thereafter, we obtain:

$$\int_{-\infty}^{\hat{\theta}} k(\theta | x_1, x_2, ..., x_n) d\theta = \int_{\hat{\theta}}^{+\infty} k(\theta | x_1, x_2, ..., x_n) d\theta = \frac{1}{2}.$$

Based on the above discussion, the procedure for the Bayesian method can be expressed in the following steps:

Step 1: Find the prior distribution, $h(\theta)$;

Step 2: Calculate $h(\theta)\sum_{i=1}^{n} f(x_i \mid \theta)$ and $\int_{-\infty}^{+\infty} h(\theta)\sum_{i=1}^{n} f(x_i \mid \theta)$, and thereafter derive the posterior distribution; and

Step 3: Minimize the posterior expected loss to find the estimator of θ.

We discuss some examples in the next section to illustrate the methods discussed above.

4. EXAMPLES

In this section, we provide examples for each of the three universal methods that have been discussed above.

4.1. Moment Method

We first provide an example of the moment method.

Example 1

Assume that $X_1, X_2, ..., X_n$ is a random sample from a population *X* with pdf given by:

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{\alpha-\beta}, & \text{if } x \in (\alpha;\beta) \\ 0, & \text{if } x \notin (\alpha;\beta) \end{cases}.$$

We wish to obtain estimates of the parameters α and β using the moment method.

Solution

Let *X* have a uniform distribution, $X \sim UNIF(0,\theta)$. Therefore:

$$E(X) = \frac{\beta+\alpha}{2}$$

$$Var(X) = E(X^2) - [E(X)]^2 = \frac{(\beta-\alpha)^2}{12}$$

$$E(X^2) = \frac{(\beta-\alpha)^2}{12} + [E(X)]^2.$$

In addition, we have: $M_1 = \overline{X}$.

By the moment method, $E(X) = M_1$, so that:

$$\frac{\beta+\alpha}{2} = \overline{X} \Leftrightarrow \beta+\alpha = 2\overline{X}. \quad (9)$$

According to the moment method, we have:

$$E(X^2) = M_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2,$$

such that:

$$\frac{(\beta-\alpha)^2}{12} + [E(X)]^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

Thereafter, we obtain:

$$(\beta-\alpha)^2 = 12\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - [E(X)]^2\right) = 12\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2\right)$$

$$= 12\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\right).$$

Hence:

$$\beta-\alpha = \sqrt{\frac{12}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}.$$

In addition, we have:

$$2\beta = 2\overline{X} + 2\sqrt{\frac{3}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}.$$

Therefore:

$$\beta = \overline{X} + \sqrt{\frac{3}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \text{ and } \alpha = \overline{X} - \sqrt{\frac{3}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}.$$

Estimation of α and β by the moment method is given by:

$$\hat{\alpha} = \overline{X} - \sqrt{\frac{3}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \text{ and } \hat{\beta} = \overline{X} + \sqrt{\frac{3}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}.$$

4.2. Maximum Likelihood Method

We now provide an example of the maximum likelihood method.

Example 2

Let $X_1, X_2, ..., X_n$ is a random sample from population *X* with pdf given by:

$$f(x,\theta) = \begin{cases} (1-\theta)x^{-\theta}, & \text{if } x \in (0,1) \\ 0, & \text{if } x \notin (0,1) \end{cases}.$$

What is the maximum likelihood estimate of θ?

Solution

The likelihood function of the random sample has the following form:

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta).$$

Thereafter, we obtain:

$$\log L(\theta) = \log\left(\prod_{i=1}^{n} f(x_i;\theta)\right) = \sum_{i=1}^{n} \log f(x_i;\theta) = \sum_{i=1}^{n} \log\left[(1-\theta)x_i^{-\theta}\right]$$
$$= n\log(1-\theta) - \theta\sum_{i=1}^{n}\log(x_i).$$

It is necessary to maximize $\log L(\theta)$ with respect to θ, as follows:

$$\frac{d\log L(\theta)}{d\theta} = \frac{d}{d\theta}\left[n\log(1-\theta) - \theta\sum_{i=1}^{n}\log(x_i)\right]$$
$$= \frac{-n}{1-\theta} - \sum_{i=1}^{n}\log(x_i).$$

Solving $\frac{d\log L(\theta)}{d\theta} = 0$ yields:

$$\frac{1}{1-\theta} = -\frac{1}{n}\sum_{i=1}^{n}\log(x_i) = -\overline{\log x} \text{ or } \theta = 1 + \frac{1}{\overline{\log x}}.$$

The maximum value of θ can be represented by examining the second derivative, so the estimate of θ is given by:

$$\hat{\theta} = 1 + \frac{1}{\overline{\log x}}.$$

4.3. Bayesian Method

Finally, we provide an example of the Bayesian method.

Example 3

Let $X_1 = 1, X_2 = 2$ be a random sample, $n = 2$, from a distribution where the pdf is described by:

$$f(x\,|\,\theta) = \binom{3}{x}\theta^x(1-\theta)^{3-x};\, x = 0,1,2,3.$$

If the prior density function of θ is:

$$h(\theta) = \begin{cases} k, & \text{if } \theta \in \left(\frac{1}{2},1\right) \\ 0, & \text{otherwise} \end{cases}.$$

What is the posterior density function of θ?

Solution

According to the assumption, we have:

$$\int_{1/2}^{1} h(\theta)d\theta = 1 \text{ and } \int_{1/2}^{1} kd\theta = 1;$$

where:

$$h(\theta)\prod_{i=1}^{n} f(x_i\,|\,\theta) = f(x_1\,|\,\theta)f(x_2\,|\,\theta)h(\theta)$$
$$= \binom{3}{x_1}\theta^{x_1}(1-\theta)^{3-x_1}\binom{3}{x_2}\theta^{x_2}(1-\theta)^{3-x_2}\,2$$
$$= 2\binom{3}{x_1}\binom{3}{x_2}\theta^{x_1+x_2}(1-\theta)^{6-x_1-x_2}$$
$$= 2\binom{3}{1}\binom{3}{2}\theta^3(1-\theta)^3 = 18\theta^3(1-\theta)^3$$

$$\int_{-\infty}^{+\infty} h(\theta)\prod_{i=1}^{n} f(x_i\,|\,\theta)d\theta = \int_{1/2}^{1} 18\theta^3(1-\theta)^3 d\theta = \frac{9}{140}.$$

The conditional distribution with sample $X_1 = 1, X_2 = 2$ is given by:

$$k(\theta\,|\,x_1 = 1, x_2 = 2) = \frac{18\theta^3(1-\theta)^3}{\frac{9}{140}} = 280\theta^3(1-\theta)^3.$$

Therefore, the posterior distribution of θ is:

$$k(\theta\,|\,x_1 = 1, x_2 = 2) = \begin{cases} 280\theta^3(1-\theta)^3, \text{ if } \theta \in \left(\frac{1}{2},1\right) \\ 0, \quad \text{otherwise} \end{cases}.$$

We discuss the advantages and disadvantages of the three universal approaches in the next section.

5. ADVANTAGES AND DISADVANTAGES OF THE UNIVERSAL APPROACHES

The three universal methods of statistical inference to determine point estimates are the moment, maximum likelihood, and Bayesian methods. These approaches can be used to calculate precise values of the parameters or moments when the data set have no missing or incomplete values. This is the primary advantage of the universal methods. In addition, these methods are easy to apply and involve simple computer algorithms. However, in practice, it is often the case that data sets are encountered with missing or incomplete values, which is the principal disadvantage of these approaches.

Cases of missing or incomplete data are a widespread issue that is frequently encountered in, for example, in the health, education and transportation fields, among others. This issue arises for many reasons, such as respondents not answering certain items in survey questions, non-acceptance and incomprehensible responses, among others (for further details, see Schafer and Graham (2002)). The issues related to missing and incomplete data can also be classified as two different types, namely missing outcomes and missing covariates.

The issues related to the estimation of parameters or moments in non-linear regression models with missing or incomplete data have been considered in a variety of topical areas, including the following: Wang *et al*. (2002) executed a JCL estimator to estimate the parameters in logistic regression with missing covariates. This method was extended by Hsieh *et al*. (2009) and Lee *et al*. (2012) in their respective research. In this vein, Lukusa *et al*. (2016) analyzed a semiparametric inverse probability weighting (SIPW) of a zero-inflated Poisson (ZIP) regression model with missing covariates.

It is clear that it is both meaningful and necessary to extend classical methods such as moment, maximum likelihood, and Bayesian methods to obtain point estimates for statistical inference to be useful when the data set contains missing or incomplete values. For further details concerning this issue, reference is made to Little (1992), Horton and Kleinman (2007), and Pho *et al*. (2018, 2019b), among others.

We discuss practical applications of the three universal approaches to Economics, Business and Decision Sciences in the next section.

6. APPLICATIONS IN ECONOMICS, BUSINESS AND DECISION SCIENCES

Estimation has been applied in many cognate fields, and is important in Economics, Finance, Business, and Decision Sciences, among other disciplines, because it can be used to determine how large-scale activities may develop, plan distributions for workers, buy materials more efficaciously, estimate project revenues, costs, and profits, and so on.

We review the applications of the methods discussed above to Economics, Business and Decision Sciences in the next sub-section.

6.1. Economics

Estimation methods have been used extensively for time series and panel data in Economics over an extended period. Sargan (1958) presented estimation methods of univariate and multivariate economic time series relationships using instrumental variables. Sargan (1961) introduced maximum likelihood estimation of economic time series relationships together with autoregressive residuals. Klein and Ozmucur (2002) provided estimation of China's economic growth rate. Chevillon and Hendry (2005) considered non-parametric direct multi-step estimation for forecasting economic processes.

Chen, Favilukis and Ludvigson (2013) presented estimation of economic models with recursive preferences. Rizal, Sahidin and Herawati (2018) introduced economic value estimation of mangrove ecosystems in Indonesia. Further interesting empirical examples in economics have been considered in Gordois *et al*. (2012), Teulings and Zubanov (2014), Hoderlein *et al*. (2017), and Cao *et al*. (2019), among others.

6.2. Business

In addition to estimation that has been widely used in Economics, estimation is also very popular in Business. Taylor and McGuire (2007) introduced a synchronous bootstrap to account for dependencies between lines of business in the estimation of the loss reserve prediction error. Anderson and Sherman (2010) applied the Fermi estimation technique to business problems. Smith (2013) developed sampling and estimation methods for use in business surveys. Obaidullah (2016) considered revisiting estimation methods of business and related tax incentives. For further interesting practical examples, see Yong and Yuanyuan (2007), Pomenkova (2010), and Casciano *et al*. (2011), among others.

6.3. Portfolio Optimization in Decision Sciences

There have also been several applications of the universal methods in Decision Sciences. The theory underlying Decision Sciences play an extremely important and significant role in the world of science and affects all aspects of decision making in the real world. A scientific theory is an explanation of a certain area of the natural world that can be tested repeatedly, using empirical, numerical or experimental data. Published scientific theories have stood up to scientific scrutiny and have contributed comprehensively to the accumulation of scientific knowledge.

The first field in decision sciences which uses the methods discussed in the paper to obtain an optimal solution is portfolio optimization. Some definitions and theory for this issue are introduced by Markowitz (1952). Pouya *et al*. (2016) solved the multi-objective portfolio optimization problem using invasive weed optimization. Fouque *et al*. (2017) developed portfolio optimization and stochastic volatility asymptotics. Olivares *et al*. (2018) provided a robust perspective on transaction costs in portfolio optimization. For further interesting practical applications, see Chen (2015), Pinto *et al*. (2015), Macedo *et al*. (2017) and Dai and Wen (2018), among others.

6.4. Bayesian Estimation

Bayesian estimation plays a very important and meaningful role in practical applications. As interesting examples, Schütt *et al*. (2016) introduced pain free and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. Dupin *et al*. (2017) provided Bayesian estimation of the global biogeographical history of the Solanaceae. Chib *et al*. (2018) presented Bayesian estimation and comparisons of moment condition models. Bernhard *et al*. (2019) analyzed Bayesian estimation of the specific shear and bulk viscosity of quark-gluon plasma. For further practical applications, see Matzke *et al*. (2015), Angelis *et al*. (2017) and Marsman *et al*. (2019), among others.

6.5. Other Disciplines

The applications in decision sciences are diverse and plentiful. There have been many research papers that have considered this issue, with the following examples. Arvai *et al*. (2004) analyzed the teaching of students to obtain superior decisions about the environment based on lessons in the decision sciences. Pidgeon and Fischhoff *et al*. (2011) introduced the role of the social and decision sciences in communicating uncertain climate risks. Chang, McAleer and Wong (2017) evaluated the connections among decision sciences, management information, and financial economics.

Chang, McAleer and Wong (2018) evaluated the connections among the decision sciences and some related cognate disciplines, such as economics, finance, business, and big data. Pho *et al*. (2019a) presented applications of the distribution functions in statistics to decision sciences. Pho *et al*. (2019c) provided applications of the optimization solution to decision sciences. Furthermore, readers may refer to the useful practical contributions of Mettler (2010), Haward and Janvier (2015), Pagell *et al*. (2019), among others.

7. CONCLUDING REMARKS

The paper presented a detailed and comprehensive approach to the theory and practical application of three universal methods in statistical inference for point estimation, namely the moment, maximum likelihood, and Bayesian methods. In addition, we discussed the advantages and dis-advantages of the three approaches, and we reviewed the practical applications of the three methods in Economics, Business and Decision Sciences.

It can be seen that the three universal approaches are popular for obtaining point estimates, and can provide precise values when the data set has no missing values. In addition to this advantage, the approaches are easy to apply and incorporate into simple algorithms. Nevertheless, in practice, we often encounter data sets that contain missing values. This is the primary disadvantage of these methods.

In the paper, we also introduced some methods to deal with the missing data problem that can be used as a combination and extension of the universal methods to analyze many important problems in the literature (see, for example, Tian and Pho (2019), Tuan *et al*. (2019), Truong *et al*. (2019), Chang *et al*. (2019) and Ly *et al.* (2019a, b), among others, for practical applications in Economics and Finance).

ACKNOWLEDGEMENTS

The authors are grateful to Phó Kim Hưng and three reviewers for very helpful comments and suggestions. For financial support, the fourth author is most grateful to the Australian Research Council and the Ministry of Science and Technology (MOST), Taiwan.

REFERENCES

Amiri, A., Allahyari, S. (2012), Change point estimation methods for control chart postsignal diagnostics: a literature review. *Quality and Reliability Engineering International*, 28(7), 673-685. https://doi.org/10.1002/qre.1266

Anderson, P.M., Sherman, C.A. (2010), Applying the Fermi estimation technique to business problems. *Journal of Applied Business and Economics*, 10(5), 33.

Angelis, K., Alvarez-Carretero, S., Dos Reis, M., Yang, Z. (2017), An evaluation of different partitioning strategies for Bayesian estimation of species divergence times. *Systematic Biology*, 67(1), 61-77. https://doi.org/10.1093/sysbio/syx061

Arvai, J.L., Campbell, V.E., Baird, A., Rivers, L. (2004), Teaching students to make better decisions about the environment: Lessons from the decision sciences. *Journal of Environmental Education*, 36(1), 33-44. https://doi.org/10.3200/JOEE.36.1.33-44

Bafumi, J., Gelman, A., Park, D.K., Kaplan, N. (2005), Practical issues in implementing and understanding Bayesian ideal point estimation. *Political Analysis*, 13(2), 171-187. https://doi.org/10.1093/pan/mpi010

Bernhard, J.E., Moreland, J.S., Bass, S.A. (2019), Bayesian estimation of the specific shear and bulk viscosity of quark–gluon plasma. *Nature Physics*, 1-5. https://doi.org/10.1038/s41567-019-0611-8

Cao, X., Lam, J.S.L. (2019), Simulation-based severe weather-induced container terminal economic loss estimation. *Maritime Policy & Management*, 46(1), 92-116. https://doi.org/10.1080/03088839.2018.1516049

Casciano, M.C., De Giorgi, V., Oropallo, F., Siesto, G. (2011), Estimation of structural business statistics for small firms by using administrative data. *Rivista di statistica ufficiale*, 13(2-3), 55-74.

Chang, C.C., Batmunkh, M.U., Wong, W.K., Jargalsaikhan, M. (2019), Relationship between capital structure and profitability: Evidence from Four Asian Tigers. *Journal of Management Information and Decision Sciences*, 22(2), 54-65. https://doi.org/10.2139/ssrn.3411977

Chang, C.L., McAleer, M., Wong, W.K. (2017), Management information, decision sciences, and financial economics: A connection. *Journal of Management Information and Decision Sciences*, 20, 1-19. https://doi.org/10.2139/ssrn.3103807

Chang, C.L., McAleer, M., Wong, W.K. (2018), Decision sciences, economics, finance, business, computing, and big data: Connections. *Advances in Decision Sciences*, 22, 1-58. https://doi.org/10.2139/ssrn.3140371

Chen, W. (2015), Artificial bee colony algorithm for constrained possibilistic portfolio optimization problem. *Physica A: Statistical Mechanics and its Applications*, 429, 125-139. https://doi.org/10.1016/j.physa.2015.02.060

Chen, X., Favilukis, J., Ludvigson, S.C. (2013), An estimation of economic models with recursive preferences. *Quantitative Economics*, 4(1), 39-83. https://doi.org/10.3982/QE97

Chevillon, G., Hendry, D.F. (2005), Non-parametric direct multi-step estimation for forecasting economic processes. *International Journal of Forecasting*, 21(2), 201-218. https://doi.org/10.1016/j.ijforecast.2004.08.004

Chib, S., Shin, M., Simoni, A. (2018), Bayesian estimation and comparison of moment condition models. *Journal of the American Statistical Association*, 113(524), 1656-1668. https://doi.org/10.1080/01621459.2017.1358172

Chung, Y., Gelman, A., Rabe-Hesketh, S., Liu, J., Dorie, V. (2015), Weakly informative prior for point estimation of covariance matrices in hierarchical models. *Journal of Educational and Behavioral Statistics*, 40(2), 136-157. https://doi.org/10.3102/1076998615570945

Dai, Z., Wen, F. (2018), Some improved sparse and stable portfolio optimization problems. *Finance Research Letters*, 27, 46-52. https://doi.org/10.1016/j.frl.2018.02.026

Dupin, J., Matzke, N.J., Särkinen, T., Knapp, S., Olmstead, R.G., Bohs, L., Smith, S.D. (2017), Bayesian estimation of the global biogeographical history of the Solanaceae. *Journal of Biogeography*, 44(4), 887-899. https://doi.org/10.1111/jbi.12898

Fouque, J.P., Sircar, R., Zariphopoulou, T. (2017), Portfolio optimization and stochastic volatility asymptotics. *Mathematical Finance*, 27(3), 704-745. https://doi.org/10.1111/mafi.12109

Goplerud, M. (2019), A Multinomial Framework for Ideal Point Estimation. *Political Analysis*, 27(1), 69-89. https://doi.org/10.1017/pan.2018.31

Gordois, A., Cutler, H., Pezzullo, L., Gordon, K., Cruess, A., Winyard, S., Chua, K. (2012), An estimation of the worldwide economic and health burden of visual impairment. *Global public health*, 7(5), 465-481. https://doi.org/10.1080/17441692.2011.634815

Harchaoui, Z., Lvy-Leduc, C. (2010), Multiple change-point estimation with a total variation penalty. *Journal of the American Statistical Association*, 105(492), 1480-1493. https://doi.org/10.1198/jasa.2010.tm09181

Haward, M.F., Janvier, A. (2015), An introduction to behavioural decision-making theories for paediatricians. *Acta Paediatrica*, 104(4), 340-345. https://doi.org/10.1111/apa.12948

Hawkins, D.M. (1976), Point estimation of the parameters of piecewise regression models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 25(1), 51-57. https://doi.org/10.2307/2346519

Hodges, J.L., Lehmann, E.L. (2012), Some problems in minimax point estimation. In Selected Works of EL Lehmann, pp. 15-30. Springer, Boston, MA. https://doi.org/10.1007/978-1-4614-1412-4_3

Hoderlein, S., Nesheim, L., Simoni, A. (2017), Semiparametric estimation of random coefficients in structural economic models. *Econometric Theory*, 33(6), 1265-1305. https://doi.org/10.1017/S0266466616000396

Horton, N.J., Kleinman, K.P. (2007), Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. *The American Statistician*, 61(1), 79-90. https://doi.org/10.1198/000313007X172556

Hsieh, S.H., Lee, S.M., Shen, P. (2009), Semiparametric analysis of randomized response data with missing covariates in logistic regression. *Computational Statistics & Data Analysis*, 53(7), 2673-2692. https://doi.org/10.1016/j.csda.2009.01.011

Klein, L.R., Ozmucur, S. (2002), The estimation of China's economic growth rate. *Journal of Economic and Social Measurement*, 28(4), 187-202. https://doi.org/10.3233/JEM-2003-0222

Lee, S.M., Li, C.S., Hsieh, S.H., Huang, L.H. (2012), Semiparametric estimation of logistic regression model with missing covariates and outcome. *Metrika*, 75(5), 621-653. https://doi.org/10.1007/s00184-011-0345-9

Lehmann, E.L. (2004), Elements of large-sample theory. Springer Science & Business Media.

Lehmann, E.L., Casella, G. (2006), Theory of point estimation. Springer Science & Business Media.

Little, R. J. (1992), Regression with missing X's: A review. *Journal of the American Statistical Association*, 87(420), 1227-1237. https://doi.org/10.2307/2290664

Lukusa, T.M., Lee, S.M., Li, C.S. (2016), Semiparametric estimation of a zero-inflated Poisson regression model with missing covariates. *Metrika*, 79(4), 457-483. https://doi.org/10.1007/s00184-015-0563-7

Ly, S., Pho, K.H., Ly, S., Wong, W.K. (2019a), Determining distribution for the quotients of dependent and independent random variables by using copulas. *Journal of Risk and Financial Management*, 12(1), 1-42. https://doi.org/10.3390/jrfm12010042

Ly, S., Pho, K.H., Ly, S., Wong, W.K. (2019b), Determining distribution for the product of random variables by using copulas. *Risks*, 7(1), 1-23. https://doi.org/10.3390/risks7010023

Macedo, L.L., Godinho, P., Alves, M.J. (2017), Mean-semivariance portfolio optimization with multiobjective evolutionary algorithms and technical analysis rules. *Expert Systems with Applications*, 79, 33-43. https://doi.org/10.1016/j.eswa.2017.02.033

Markowitz, H. (1952), Portfolio selection. *Journal of Finance*, *7*(1), 77-91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

Marsman, M., Waldorp, L., Dablander, F., Wagenmakers, E.J. (2019), Bayesian estimation of explained variance in ANOVA designs. *Statistica Neerlandica*, 73(3), 351-372. https://doi.org/10.1111/stan.12173

Matzke, D., Dolan, C.V., Batchelder, W.H., Wagenmakers, E.J. (2015), Bayesian estimation of multinomial processing tree models with heterogeneity in participants and items. *Psychometrika*, 80(1), 205-235. https://doi.org/10.1007/s11336-013-9374-9

Mettler, T. (2010), Thinking in terms of design decisions when developing maturity models. *International Journal of Strategic Decision Sciences*, 1(4), 76-87. https://doi.org/10.4018/jsds.2010100105

Obaidullah, M. (2016). Revisiting estimation methods of business zakat and related tax incentives. *Journal of Islamic Accounting and Business Research*, 7(4), 349-364. https://doi.org/10.1108/JIABR-10-2014-0035

Olivares-Nadal, A.V., DeMiguel, V. (2018), A robust perspective on transaction costs in portfolio optimization. *Operations Research*, 66(3), 733-739. https://doi.org/10.1287/opre.2017.1699

Pagell, M., Wiengarten, F., Fan, D., Humphreys, P., Lo, C.K. (2019), Managerial time horizons and the decision to put operational workers at risk: The role of debt. *Decision Sciences*, 50(3), 582-611. https://doi.org/10.1111/deci.12338

Pearson, K. (1902), On the systematic fitting of curves to observations and measurements. *Biometrika*, 1(3), 265-303. https://doi.org/10.2307/2331540

Pidgeon, N., Fischhoff, B. (2011), The role of social and decision sciences in communicating uncertain climate risks. *Nature Climate Change*, 1(1), 35. https://doi.org/10.1038/nclimate1080

Pinto, T., Morais, H., Sousa, T.M., Sousa, T., Vale, Z., Praca, I., Pires, E.J.S. (2015), Adaptive portfolio optimization for multiple electricity markets participation. *IEEE Transactions on Neural Networks and Learning Systems*, 27(8), 1720-1733. https://doi.org/10.1109/TNNLS.2015.2461491

Pho, K.H., Nguyen, V.T. (2018), Comparison of Newton-Raphson algorithm and Maxlik function. *Journal of Advanced Engineering and Computation*, 2(4), 281-292. https://doi.org/10.25073/jaec.201824.219

Pho, K.H., Ho, T.D.C., Tran, T.K., Wong, W.K. (2019a), Moment generating function, expectation and variance of ubiquitous distributions with applications in decision sciences: A review. *Advances in Decision Sciences*, 23(2), 1-85. https://doi.org/10.2139/ssrn.3430778

Pho, K.H., Ly, S., Ly, S., Lukusa, T.M. (2019b), Comparison among Akaike information criterion, Bayesian information criterion and Vuong's test in model selection: A case study of violated speed regulation in Taiwan. *Journal of Advanced Engineering and Computation*, 3(1), 293-303. https://doi.org/10.25073/jaec.201931.220

Pho, K.H., Tran, T.K., Ho, T.D.C., Wong, W.K. (2019c), Optimal solution techniques in decision sciences: A review. *Advances in Decision Sciences* 23(1), 1-47. https://doi.org/10.2139/ssrn.3430764

Pomenkova, J. (2010), An alternative approach to the dating of business cycle: Nonparametric kernel estimation. *Prague Economic Papers*, 19(3), 251-272. https://doi.org/10.18267/j.pep.375

Pouya, A.R., Solimanpur, M., Rezaee, M. J. (2016), Solving multi-objective portfolio optimization problem using invasive weed optimization. *Swarm and Evolutionary Computation*, 28, 42-57. https://doi.org/10.1016/j.swevo.2016.01.001

Rizal, A., Sahidin, A., Herawati, H. (2018), Economic value estimation of mangrove ecosystems in Indonesia. *Biodiversity International Journal*, 2(1), 98-100. https://doi.org/10.15406/bij.2018.02.00051

Sargan, J.D. (1958), The estimation of economic relationships using instrumental variables. *Econometrica*, 393-415. https://doi.org/10.2307/1907619

Sargan, J.D. (1961), The maximum likelihood estimation of economic relationships with autoregressive residuals. *Econometrica*, 414-426. https://doi.org/10.2307/1909642

Schafer, J.L., Graham, J.W. (2002), Missing data: our view of the state of the art. *Psychological Methods*, 7(2), 147. https://doi.org/10.1037/1082-989X.7.2.147

Schennach, S.M. (2007), Point estimation with exponentially tilted empirical likelihood. *Annals of Statistics*, 35(2), 634-672. https://doi.org/10.1214/009053606000001208

Schütt, H.H., Harmeling, S., Macke, J.H., Wichmann, F.A. (2016), Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Research*, 122, 105-123. https://doi.org/10.1016/j.visres.2016.02.002

Smith, P. (2013), Sampling and estimation for business surveys. *Designing and conducting business surveys*, 165-218. https://doi.org/10.1002/9781118447895.ch05

Taylor, G., McGuire, G. (2007), A synchronous bootstrap to account for dependencies between lines of business in the estimation of loss reserve prediction error. *North American Actuarial Journal*, 11(3), 70-88. https://doi.org/10.1080/10920277.2007.10597467

Teulings, C.N., Zubanov, N. (2014), Is economic recovery a myth? Robust estimation of impulse responses. *Journal of Applied Econometrics*, 29(3), 497-514. https://doi.org/10.1002/jae.2333

Tian. Y, Pho, K.H. (2019), A statistical view to study the aphorisms in Nahj al Balaghah. *Digital Scholarship in the Humanities*. (In Press). https://doi.org/10.1093/llc/fqz075

Truong, B.C., Pho, K.H., Nguyen, V.B., Tuan, B.A., Wong, W.K. (2019), Graph theory and environmental algorithmic solutions to assign vehicles: Application to garbage collection in Vietnam. *Advances in Decision Sciences*, 23(2), 1-35. https://doi.org/10.2139/ssrn.3430459

Tuan, B.A., Pudikova, G.N., Mahmoudi, M.R., Pho, K.H. (2019), Statistical approaches in literature: Comparing and clustering the alternatives of love in Divan of Hafiz. Digital Scholarship in the Humanities. (In Press). https://doi.org/10.1093/llc/fqz069

Wang, C.Y., Chen, J.C., Lee, S.M., Ou, S.T. (2002), Joint conditional likelihood estimator in logistic regression with missing covariate data. *Statistica Sinica*, 555-574.

Yong, Z., Yuanyuan, L. (2007), The estimation of technological progress on the energy consumption returns effects. *Economist*, 2, 45-52.

Yoo, J.H., Lee, S.W., Park, S.K., Kim, D.H. (2017), A robust lane detection method based on vanishing point estimation using the relevance of line segments. *IEEE Transactions on Intelligent Transportation Systems*, 18(12), 3254-3266. https://doi.org/10.1109/TITS.2017.2679222

Received on 08-11-2019 Accepted on 27-11-2019 Published on 13-12-2019

DOI: https://doi.org/10.6000/1929-7092.2019.08.90