E-Bayesian Estimation and Hierarchical Bayesian Estimation of Poisson Distribution Parameter under Entropy Loss Function

Chunping Li, Huibing Hao

Abstract—**In this paper, a new parameter estimation method, named E-Bayesian estimation, is introduced to estimate the unknown parameter of the Poisson distribution. Based on the entropy loss function, formulas of E-Bayesian estimation and hierarchical Bayesian estimation for the unknown parameter are given, and the properties of E-Bayesian estimation and hierarchical Bayesian estimation are also discussed. Compare with hierarchical Bayesian estimation, the expression of E-Bayesian estimation is simpler; therefore, the E-Bayesian estimation method is worth being recommended. Finally, a real data of air crash happened in China is used to demonstrate the usefulness and validity of the proposed model and method.**

Index Terms—**Entropy loss function, E-Bayesian estimation, hierarchical Bayesian estimation, Poisson distribution**

I. INTRODUCTION

IN practical applications, the Poisson distribution model is often used to describe successfully number, such as the number of birth defects, the number of machine failures, the number of natural disasters, the number of waiting for guests on the bus platform, and so on. For Poisson distribution, many different parameter estimation methods have been presented, such as the maximum likelihood estimation[1], and the Bayesian estimation[2].

The idea of hierarchical Bayesian estimation was first introduced by Lindley and Smith[3], and many researchers[4-6] have used a hierarchical Bayesian method to obtain the estimation of the unknown parameter. However, this method encountered complicated integration, and the complicated integration was very hard to be implemented in practice. In this situation, a new estimation method, named E-Bayesian estimation, was introduced by Han[5]. Compare with the hierarchical Bayesian method, the E-Bayesian estimation method is simpler.

In recent years, there has been a growing interest in the study of E-Bayesian and hierarchical Bayesian estimation under the different distributions, such as exponential distribution[5], binomial distribution[6], Pascal distribution[7-8], and Pareto distribution[9].

Manuscript received August 2, 2018. This work was supported by the Humanity and Social Science Foundation for the Ministry of Education of China (No. 19YJAZH039), the Key Project of Hubei Provincial Education Department (No.D20172701), the National Bureau of Statistics of China (No.2017LY73), the Technology Creative Project of Excellent Middle & Young Team of Hubei Province (2019).

C. P. Li is with the Department of Mathematics, Hubei Engineering University, Hubei, 432000, China. e-mail: lichunping315@163.com.

[a] H. B. Hao is the corresponding author with the Department of Mathematics, Hubei Engineering University, Hubei, 432000, China. e-mail: haohuibing1979@163.com.

[b]H. B. Hao is with the Hubei Key Laboratory of Applied Mathematics, Faculty of Mathematics and Statistics, Hubei University, Wuhan, 430062, China. e-mail: haohuibing1979@163.com.

Although a lot of work has been done on the statistical inferences of the unknown parameters of the Poisson distribution, however, E-Bayesian and hierarchical Bayesian estimations of Poisson distribution under entropy loss function have not been addressed so for.

In this paper, E-Bayesian and Hierarchical Bayesian estimations of unknown parameters of the Poisson distribution are presented under entropy loss function. The rest of paper is organized as follow: in section II, the Poisson distribution and the definition of entropy loss function are introduced. In section III, the definition of E-Bayesian estimation is described. In section VI to VI, the formulas of the Bayesian estimation, E-Bayesian estimation and hierarchical Bayesian estimation are obtained. In section VII, the properties of E-Bayesian estimation and hierarchical Bayesian estimation are discussed. In section VIII, a real-data set example is given. Section IX is the conclusion.

II. POISSON DISTRIBUTION AND ENTROPY LOSS FUNCTION

Suppose that Poisson distribution $P(\lambda)$ has the following distribution law

$$P(X=x)=\frac{\lambda^x}{x!}e^{-\lambda}, x=0,1,\cdots. \tag{1}$$

where λ ($\lambda>0$) is unknown parameter. And suppose x = ($x_1, \dots, x_n$) is a random observation sample from $P(\lambda)$, then, the likelihood function of the observed sample can be obtained as

$$L(x\mid\lambda)=\prod_{i=1}^{n}P(x_i\,|\lambda)=\frac{\lambda^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i}e^{-n\lambda} \tag{2}$$

In order to obtain the Bayesian estimation, most of Bayesian inference procedures have been developed under the square error loss function or Linex loss function. In many practical situations, it appears to be more realistic to express the loss in terms of the ratio, and the entropy loss function has this property. The entropy loss function is first introduced by Calabria and Pulcini [10].

Definition 1 If random variable X_i has density function $f(x_i;\theta)$, the entropy loss function can be defined as

$$L(\theta,\delta)=E\left\{\ln\frac{f(\theta,x_1,x_2,\cdots,x_n)}{f(\delta,x_1,x_2,\cdots,x_n)}\right\} \tag{3}$$

where δ is an estimator of parameter θ .

(Advance online publication: 12 August 2019)

Using $P(\lambda)$ distribution in Equation (1) and its likelihood function in Equation (2), the entropy loss function can be given as

$$L(\lambda,\delta)=E\left\{\ln\frac{p(\lambda,x_1,x_2,\cdots,x_n)}{p(\delta,x_1,x_2,\cdots,x_n)}\right\}$$

$$=E\left[\ln\frac{\lambda^{\sum_{i=1}^{n}x_i}e^{-n\lambda}/\prod_{i=1}^{n}x_i}{\delta^{\sum_{i=1}^{n}x_i}e^{-n\delta}/\prod_{i=1}^{n}x_i}\right]=n\lambda(\ln\frac{\lambda}{\delta}+\frac{\delta}{\lambda}-1) \quad (4)$$

Obviously, $L(\lambda,\delta)$ is strictly convex function of λ.

III. Definition of E-Bayesian estimation

In this section, we consider Bayesian estimation of the parameter λ under the entropy loss function, and we assume that $\lambda\sim Gamma(a,b)$ has gamma prior distribution as follow

$$\pi(\lambda\mid a,b)=\frac{b^a\lambda^{a-1}}{\Gamma(a)}\exp\left(-b\lambda\right) \quad (5)$$

where $\Gamma(a)=\int_0^{+\infty}x^{a-1}\exp(-x)dx$ is a gamma function, and hyper parameters $a>0$ and $b>0$.

According to Han [5], the selection of the hyper parameters a and b should guarantee that $\pi(\lambda|a,b)$ is a decreasing function of λ. The derivative of $\pi(\lambda|a,b)$ with respect to λ is

$$\frac{d[\pi(\lambda\mid a,b)]}{d\lambda}=\frac{b^a\lambda^{a-2}\exp(-b\lambda)}{\Gamma(a)}[(a-1)-b\lambda] \quad (6)$$

In Equation (5), when $0<a<1$, $d[\pi(\lambda\mid a,b)]/d\lambda<0$, then, $\pi(\lambda| a, b)$ is a decreasing function of λ. Given $0<a<1$, if the b is larger, then the tail of the Gamma density function will be thinner. Considering the robustness of Bayesian estimation, the thinner tailed prior distribution can reduce the robustness of Bayesian estimation. For this reason, b should not be larger than c, where c ($c>0$) is an upper bound. Therefore, the hyper-parameters a and b should be selected with the restriction of $0<a<1$ and $0<b<c$. How to determine the constant c would be based on expert's opinion.

The definition for E-Bayesian estimation was originally addressed by Han [5] as follow.

Definition 2 Suppose $\hat{\lambda}_B$ is the Bayesian estimation of λ, and $\pi(a, b)$ is the prior density function of the hyper parameters a and b, then the E-Bayesian estimation of λ (expectation of the Bayesian estimate of λ) can be given by $\hat{\lambda}_{EB}$ as follow

$$\hat{\lambda}_{EB}=\iint_D\hat{\lambda}_B\left(a,b\right)\pi\left(a,b\right)dadb=E[\hat{\lambda}_B\left(a,b\right)] \quad (7)$$

where D is the domain of a and b.

Definition 2 indicates that the E-Bayesian estimation of λ is just the expectation of the Bayesian estimation of λ for all the hyper-parameters a and b.

IV. Bayesian estimation under entropy loss function

Suppose $x=(x_1,\dots,x_n)$ is a random sample from $P(\lambda)$, then, the posterior density function of λ can be written as

$$\pi(\lambda\mid x)=\frac{L(x\mid\lambda)\pi(\lambda\mid a,b)}{\int_0^{\infty}L(x\mid\lambda)\pi(\lambda\mid a,b)d\lambda}$$

$$=\frac{(b+n)^{a+T}}{\Gamma(a+T)}\lambda^{a+T-1}\exp\left(-(b+n)\lambda\right) \quad (8)$$

where $T=\sum_{i=1}^{n}x_i$.

Theorem 1 Suppose the entropy loss function is given by Equation (4), for any prior distribution $\pi(\lambda)$ of λ, if there is a estimator δ_B, and the corresponding Bayes risk $R(\delta_B)<+\infty$, then we can obtain a unique Bayesian estimation of λ as

$$\delta_B=E\left[\lambda|x\right] \quad (9)$$

Proof Let δ_B be any Bayesian estimation of λ, under the entropy loss function in Equation (4), the corresponding Bayes risk $R(\delta_B)$ can be obtained as

$$R(\delta_B)=E\left[L(\lambda,\delta_B)\right]=nE\left[\lambda(\ln\frac{\lambda}{\delta_B}+\frac{\delta_B}{\lambda}-1)\right]$$

$$=nE\left\{E\left[\lambda(\ln\frac{\lambda}{\delta_B}+\frac{\delta_B}{\lambda}-1)|x\right]\right\}$$

In order to minimize $R(\delta_B)$, let

$$g(\delta_B)=E\left\{E\left[\lambda(\ln\frac{\lambda}{\delta_B}+\frac{\delta_B}{\lambda}-1)|x\right]\right\}$$

$$=E\left[\lambda\ln\lambda|x\right]-\ln\delta_B E\left[\lambda|x\right]+\delta_B-E\left[\lambda|x\right]$$

Taking the derivative of $g(\delta_B)$ with respect to δ_B, and solving the equation $d[g(\delta_B)]/d\delta_B=0$, we can obtain the equation's solution as

$$\delta_B=E\left[\lambda|x\right] \quad (10)$$

Next, we will proof the uniqueness of δ_B. Considering that $g(\delta_B)$ is strictly convex function of δ_B, then we know that $\delta_B=E\left[\lambda|x\right]$ is a unique Bayesian estimation of λ.

Theorem 2 If $x=(x_1,\dots, x_n)$ is a random observation sample from $P(\lambda)$, the likelihood function is given by Equation (2), the prior density distribution $\pi(\alpha| a, b)$ of λ is given by Equation (5), under the entropy loss function, then Bayesian estimation of λ can be obtained as

$$\hat{\lambda}_B=\frac{a+T}{b+n} \quad (11)$$

And the estimate $\hat{\lambda}_B$ is admissible.

Proof From $\lambda|x\sim Gamma(a+T,n+b)$, we can get

$$E\left(\lambda\mid x\right)=\frac{a+T}{n+b}$$

Based on Theorem 1, we can get the unique Bayesian estimation of λ as

$$\hat{\lambda}_B=E\left(\lambda\mid x\right)=\frac{a+T}{n+b}$$

Obviously, the estimate $\hat{\lambda}_B$ is admissible.

V. E-Bayesian estimation under entropy loss function

In this section, we will introduce the E-Bayesian estimations of λ under the entropy loss function, and using three different prior distributions of the hyper parameters a and b.

Theorem 3 Under the entropy loss function is given by Equation (4), suppose $x=(x_1, \dots, x_n)$ is a random observation sample from $P(\lambda)$,let $T=\sum_{i=1}^{n} x_i$ and λ has gamma prior distribution as shown in Equation (5), we have

(1) If the prior density function of hyper parameters (a, b) is

$$\pi_1(a,b)=\frac{1}{c}, 0<a<1, 0<b<c \tag{12}$$

then, the corresponding E-Bayesian estimation of λ is

$$\hat{\lambda}_{EB1}=\frac{(1+2T)}{2c}\ln\left(\frac{n+c}{n}\right) \tag{13}$$

(2) If the prior density function of hyper parameters (a, b) is

$$\pi_2(a,b)=\frac{2(c-b)}{c^2}, 0<a<1, 0<b<c \tag{14}$$

then, the corresponding E-Bayesian estimation of λ is

$$\hat{\lambda}_{EB2}=\frac{(1+2T)}{c^2}\left[(c+n)\ln\left(\frac{n+c}{n}\right)-c\right] \tag{15}$$

(3) If the prior density function of hyper parameters (a, b) is

$$\pi_3(a,b)=\frac{2b}{c^2}, 0<a<1, 0<b<c \tag{16}$$

then, the corresponding E-Bayesian estimation of λ is

$$\hat{\lambda}_{EB3}=\frac{(1+2T)}{c^2}\left[c-n\ln\left(\frac{n+c}{n}\right)\right] \tag{17}$$

Proof (1) From the definition of E-Bayesian estimation in Equation (7) and the Bayesian estimation in Equation (9), we can get

$$\hat{\lambda}_{EB1}=\iint_D \hat{\lambda}_B \pi_1(a,b)\,dadb$$
$$=\frac{1}{c}\int_0^c\int_0^1 \frac{a+T}{b+n}dadb=\frac{(1+2T)}{2c}\ln\left(\frac{n+c}{n}\right)$$

(2) Similarly, the E-Bayesian estimation of λ based on π_2(a, b) can be obtained as

$$\hat{\lambda}_{EB2}=\iint_D \hat{\lambda}_B \pi_2(a,b)\,dadb$$
$$=\frac{2}{c^2}\int_0^c\int_0^1 (c-b)\frac{(a+T)}{b+n}dadb$$
$$=\frac{(1+2T)}{c^2}\left[(c+n)\ln\left(\frac{n+c}{n}\right)-c\right]$$

(3) Similarly, the E-Bayesian estimation of λ based on π_3(a, b) can be obtained as

$$\hat{\lambda}_{EB3}=\iint_D \hat{\lambda}_B \pi_3(a,b)\,dadb$$
$$=\frac{2}{c^2}\int_0^c\int_0^1 b\frac{(a+T)}{b+n}dadb$$
$$=\frac{(1+2T)}{c^2}\left[c-n\ln\left(\frac{n+c}{n}\right)\right]$$

VI. Hierarchical Bayesian estimation under entropy loss function

In this section, we will derive the hierarchical Bayesian estimation of the parameter λ under entropy loss function. The definition for hierarchical Bayesian estimation was originally addressed by Lindley and Smith [3]. Suppose the function law of random variable is $P(x;\lambda)$, $\pi(\lambda|a, b)$ is the prior density function of the parameter λ, $\pi(a, b)$ is the prior density function of the hyper parameters a and b, then the corresponding hierarchical prior density functions of λ can be expressed as

$$\pi(\lambda)=\iint_D \pi(\lambda|a,b)\pi(a,b)\,dadb \tag{18}$$

Theorem 4 Under the entropy loss function in Equation (4), suppose $x=(x_1, \dots, x_n)$ is a random observation sample from $P(\lambda)$, let $T=\sum_{i=1}^{n} x_i$ and λ has gamma prior distribution in Equation (15), we can get

(1) If the prior density function of hyper parameters (a, b) is given by Equation (12), the corresponding hierarchical Bayesian estimation of λ is

$$\hat{\lambda}_{HB1}=\frac{\int_0^c\int_0^1 \frac{\Gamma(T+a+1)}{\Gamma(a)}\frac{b^a}{(b+n)^{1+a+T}}dadb}{\int_0^c\int_0^1 \frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^a}{(b+n)^{a+T}}dadb} \tag{19}$$

(2) If the prior density function of hyper parameters (a, b) is given by Equation (14), the corresponding hierarchical Bayesian estimation of λ is

$$\hat{\lambda}_{HB2}=\frac{\int_0^c\int_0^1 (c-b)\frac{\Gamma(T+a+1)}{\Gamma(a)}\frac{b^a}{(b+n)^{(T+a+1)}}dadb}{\int_0^c\int_0^1 (c-b)\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^a}{(b+n)^{(T+a)}}dadb} \tag{20}$$

(3) If the prior density function of hyper parameters (a, b) is given by Equation (16), the corresponding hierarchical Bayesian estimation of λ is

$$\hat{\lambda}_{HB3}=\frac{\int_0^c\int_0^1 \frac{\Gamma(T+a+1)}{\Gamma(a)}\frac{b^{a+1}}{(b+n)^{(T+a+1)}}dadb}{\int_0^c\int_0^1 \frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^{a+1}}{(b+n)^{(T+a)}}dadb} \tag{21}$$

Proof (1) From the prior density function $\pi(\lambda|a, b)$ of parameter λ, and the prior density function $\pi_1(a,b)$ of hyper parameters (a, b) , we can get the hierarchical prior density function of λ as

$$\pi_1(\lambda)=\iint_D \pi(\lambda|a,b)\pi(a,b)\,dadb=\frac{1}{c}\int_0^c\int_0^1 \frac{b^a\lambda^{a-1}}{\Gamma(a)}\exp(-b\lambda)\,dadb$$

Then, the hierarchical posterior density function of λ can be given as

$$h_1(\lambda|x)=\frac{L(x|\lambda)\pi_1(\lambda)}{\int_0^{+\infty} L(x|\lambda)\pi_1(\lambda)\,d\lambda}$$
$$=\frac{\int_0^c\int_0^1 \frac{b^a\lambda^{T+a-1}\exp(-(b+n)\lambda)}{\Gamma(a)}dadb}{\int_0^{+\infty}\int_0^c\int_0^1 \frac{b^a\lambda^{T+a-1}\exp(-(b+n)\lambda)}{\Gamma(a)}dadbd\lambda}$$
$$=\frac{\int_0^c\int_0^1 \frac{b^a\lambda^{a+T-1}e^{-(n+b)\lambda}}{\Gamma(a)}dadb}{\int_0^c\int_0^1 \frac{\Gamma(T+a)b^a}{\Gamma(a)(b+n)^{T+a}}dadb}$$

Then, the hierarchical Bayesian estimation of λ can be given as

$$\hat{\lambda}_{HB1}=E[\lambda|x]$$

$$=\frac{\int_0^{+\infty}\int_0^c\int_0^1\frac{b^a\lambda^{a+T}\exp(-(n+b)\lambda)}{\Gamma(a)}dadbd\lambda}{\int_0^c\int_0^1\frac{\Gamma(T+a)b^a}{\Gamma(a)(b+n)^{T+a}}dadb}$$

$$=\frac{\int_0^c\int_0^1\frac{\Gamma(T+a+1)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a+1}}dadb}{\int_0^c\int_0^1\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a}}dadb}$$

(2) From the prior density function of parameter λ, and the prior density function of hyper parameters (a, b)

$$\pi(a,b)=\frac{2(c-b)}{c^2},0<a<1,0<b<c$$

we can get the hierarchical prior density functions of λ as

$$\pi_2(\lambda)=\iint_D\pi(\lambda/a,b)\pi(a,b)dadb$$

$$=\frac{2}{c^2}\int_0^c\int_0^1(c-b)\frac{b^a\lambda^{a-1}}{\Gamma(a)}\exp(-b\lambda)dadb$$

Then, the hierarchical posterior density functions of λ can be given as

$$h_2(\lambda\,|\,x)=\frac{L(x\,|\,\lambda)\pi_2(\lambda)}{\int_0^{+\infty}L(x\,|\,\lambda)\pi_2(\lambda)d\lambda}$$

$$=\frac{\int_0^c\int_0^1(c-b)\frac{b^a\lambda^{T+a-1}\exp(-(b+n)\lambda)}{\Gamma(a)}dadb}{\int_0^{+\infty}\int_0^c\int_0^1(c-b)\frac{b^a\lambda^{T+a-1}\exp(-(b+n)\lambda)}{\Gamma(a)}dadbd\lambda}$$

$$=\frac{\int_0^c\int_0^1(c-b)\frac{b^a\lambda^{T+a-1}\exp(-(b+nT)\lambda)}{\Gamma(a)}dadb}{\int_0^c\int_0^1(c-b)\frac{\Gamma(T+a)b^a}{\Gamma(a)(b+T)^{T+a}}dadb}$$

Then, the hierarchical Bayesian estimation of λ can be given as

$$\hat{\lambda}_{HB2}=E[\lambda|x]$$

$$=\frac{\int_0^{+\infty}\int_0^c\int_0^1\lambda(c-b)\frac{b^a\lambda^{T+a-1}\exp(-(b+n)\lambda)}{\Gamma(a)}dadbd\lambda}{\int_0^c\int_0^1(c-b)\frac{\Gamma(T+a)b^a}{\Gamma(a)(b+n)^{T+a}}dadb}$$

$$=\frac{\int_0^c\int_0^1(c-b)\frac{\Gamma(T+a+1)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a+1}}dadb}{\int_0^c\int_0^1(c-b)\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a}}dadb}$$

(3) From the prior density function of parameter λ, and the prior density function of hyper parameters (a, b)

$$\pi(a,b)=\frac{2b}{c^2},0<a<1,0<b<c$$

we can get the hierarchical prior density functions of λ as

$$\pi_3(\lambda)=\iint_D\pi(\lambda/a,b)\pi(a,b)dadb$$

$$=\frac{2}{c^2}\int_0^c\int_0^1\frac{b^{a+1}\lambda^{a-1}}{\Gamma(a)}\exp(-b\lambda)dadb$$

Then, the hierarchical posterior density functions of λ can be given as

$$h_3(\lambda\,|\,x)=\frac{L(x\,|\,\lambda)\pi_3(\lambda)}{\int_0^{+\infty}L(x\,|\,\lambda)\pi_3(\lambda)d\lambda}$$

$$=\frac{\int_0^c\int_0^1\frac{b^{a+1}\lambda^{T+a-1}\exp(-(b+n)\lambda)}{\Gamma(a)}dadb}{\int_0^{+\infty}\int_0^c\int_0^1\frac{b^{a+1}\lambda^{T+a-1}\exp(-(b+n)\lambda)}{\Gamma(a)}dadbd\lambda}$$

$$=\frac{\int_0^c\int_0^1\frac{b^{a+1}\lambda^{T+a-1}\exp(-(b+n)\lambda)}{\Gamma(a)}dadb}{\int_0^c\int_0^1\frac{\Gamma(T+a)b^{a+1}}{\Gamma(a)(b+n)^{T+a}}dadb}$$

Then, the hierarchical Bayesian estimation of λ can be given as

$$\hat{\lambda}_{HB3}=E[\lambda|x]$$

$$=\frac{\int_0^{+\infty}\int_0^c\int_0^1\lambda\frac{b^{a+1}\lambda^{T+a-1}\exp(-(b+n)\lambda)}{\Gamma(a)}dadbd\lambda}{\int_0^c\int_0^1\frac{\Gamma(T+a)b^{a+1}}{\Gamma(a)(b+n)^{T+a}}dadb}$$

$$=\frac{\int_0^c\int_0^1\frac{\Gamma(T+a+1)}{\Gamma(a)}\frac{b^{a+1}}{(b+n)^{T+a+1}}dadb}{\int_0^c\int_0^1\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^{a+1}}{(b+n)^{T+a}}dadb}$$

VII. COMPARE E-BAYESIAN ESTIMATION AND HIERARCHICAL BAYESIAN ESTIMATION

In this section, we will discuss the relationship of E-Bayesian and hierarchical Bayesian estimations under the entropy loss function.

Theorem 5 In Theorems 3 and 4, for $\hat{\lambda}_{EBi}$ and $\hat{\lambda}_{HBi}$, $i=1,2,3$, then we have

$$\lim_{n\to\infty}\hat{\lambda}_{EB1}=\lim_{n\to\infty}\hat{\lambda}_{HB1}=0,$$
$$\lim_{n\to\infty}\hat{\lambda}_{EB2}=\lim_{n\to\infty}\hat{\lambda}_{HB2}=0,$$
$$\lim_{n\to\infty}\hat{\lambda}_{EB3}=\lim_{n\to\infty}\hat{\lambda}_{HB3}=0 \tag{22}$$

Proof (1) From the Theorem 4 and the properties of Gamma function, we have

$$\hat{\lambda}_{HB1}=\frac{\int_0^c\int_0^1\frac{\Gamma(T+a+1)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a+1}}dadb}{\int_0^c\int_0^1\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a}}dadb}$$

$$=\frac{\int_0^c\int_0^1\frac{T+a}{b+n}\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a}}dadb}{\int_0^c\int_0^1\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a}}dadb}$$

When $0<a<1, 0<b<c$, we know that $\frac{T+a}{b+n}$ is continuous, and $\frac{\Gamma(T+a)b^a}{\Gamma(a)(b+n)^{T+a}}>0$, By using the mean value theorem of integral, there is at least one number $a_1\in(0,1)$ and $b_1\in(0,c)$, we have

$$\int_0^c\int_0^1\frac{(T+a)}{(b+n)}\frac{\Gamma(T+a)b^a}{\Gamma(a)(b+n)^{T+a}}dadb$$
$$=\frac{(T+a_1)}{(b_1+n)}\int_0^c\int_0^1\frac{\Gamma(T+a)b^a}{\Gamma(a)(b+n)^{T+a}}dadb$$

then

$$\lim_{n\to+\infty}\hat{\lambda}_{HB1}=\lim_{n\to+\infty}\frac{\int_0^c\int_0^1\frac{\Gamma(T+a+1)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a+1}}dadb}{\int_0^c\int_0^1\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a}}dadb}$$
$$=\lim_{n\to+\infty}\frac{T+a_1}{b_1+n}=0$$

According to Equation (13), we can get

$$\hat{\lambda}_{EB1}=\frac{(2T+1)}{2c}\ln\left(\frac{n+c}{n}\right)$$

Taking the limit of both sides, obviously we can get

$$\lim_{n\to+\infty}\hat{\lambda}_{EB1}=\lim_{n\to+\infty}\frac{(2T+1)}{2c}\ln\left(\frac{n+c}{n}\right)=0$$

Then

$$\lim_{n\to\infty}\hat{\lambda}_{EB1}=\lim_{n\to\infty}\hat{\lambda}_{HB1}=0$$

2) Similarly, there is at least one number $a_2\in(0,1)$ and $b_2\in(0,c)$, we have

$$\hat{\lambda}_{HB2}=\frac{T+a_2}{b_2+n}\frac{\int_0^c\int_0^1(c-b)\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a}}dadb}{\int_0^c\int_0^1(c-b)\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a}}dadb}=\frac{T+a_2}{b_2+n}$$

Taking the limit of both sides, we get

$$\lim_{n\to+\infty}\hat{\lambda}_{HB2}=\lim_{n\to+\infty}\frac{\int_0^c\int_0^1(c-b)\frac{\Gamma(T+a+1)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a+1}}dadb}{\int_0^c\int_0^1(c-b)\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^a}{(b+n)^{T+a}}dadb}$$
$$=\lim_{n\to+\infty}\frac{T+a_2}{b_2+n}=0$$

According to Equation (15), we can get

$$\hat{\lambda}_{EB2}=\frac{2T+1}{c^2}\left[(c+n)\ln\left(\frac{n+c}{n}\right)-c\right]$$

Taking the limit of both sides, we get

$$\lim_{n\to+\infty}\hat{\lambda}_{EB2}=\lim_{n\to+\infty}\frac{2T+1}{c^2}\left[(c+n)\ln\left(\frac{n+c}{n}\right)-c\right]=0$$

Then

$$\lim_{n\to\infty}\hat{\lambda}_{EB2}=\lim_{n\to\infty}\hat{\lambda}_{HB2}=0$$

3) Similarly, there is at least one number $a_3\in(0,1)$ and $b_3\in(0,c)$, we have

$$\hat{\lambda}_{HB3}=\frac{\int_0^c\int_0^1\frac{T+a}{b+n}\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^{a+1}}{(b+n)^{T+a}}dadb}{\int_0^c\int_0^1\frac{\Gamma(n+a)}{\Gamma(a)}\frac{b^{a+1}}{(b+n)^{T+a}}dadb}=\frac{T+a_3}{b_3+n}$$

Taking the limit of both sides, we get

$$\lim_{n\to+\infty}\hat{\lambda}_{HB3}=\lim_{n\to+\infty}\frac{\int_0^c\int_0^1\frac{T+a}{b+n}\frac{\Gamma(T+a)}{\Gamma(a)}\frac{b^{a+1}}{(b+n)^{T+a}}dadb}{\int_0^c\int_0^1\frac{\Gamma(n+a)}{\Gamma(a)}\frac{b^{a+1}}{(b+n)^{T+a}}dadb}$$
$$=\lim_{n\to+\infty}\frac{T+a_3}{b_3+n}=0$$

According to Equation (17), we get

$$\hat{\lambda}_{EB3}=\frac{2T+1}{c^2}\left[c-n\ln\left(\frac{n+c}{n}\right)\right]$$

Taking the limit of both sides, we get

$$\lim_{n\to+\infty}\hat{\lambda}_{EB3}=\lim_{n\to+\infty}\frac{2T+1}{c^2}\left[c-n\ln\left(\frac{n+c}{n}\right)\right]=0$$

Then

$$\lim_{n\to\infty}\hat{\lambda}_{EB3}=\lim_{n\to\infty}\hat{\lambda}_{HB3}=0$$

VIII. A REAL EXAMPLE

To illustrate the proposed model and method in this paper, a real data of air crash happened in China in recent 30 years (1981-2010) as shown in Table I, and the data follows a Poisson distribution. In this data set, *k* represents the number of air crash happened in 30 years, n_k represents the frequency of occurrence of air crash.

TABLE I
THE DATA OF AIR CRASH HAPPENED

k	0	1	2	3	4	5	total
n_k	13	10	3	3	0	1	30

From Theorems 2-5 and Table I, we can get the estimation of λ_{EBi} and λ_{HBi} (i=1, 2, 3), and some numerical results are listed in Table II

From Table II, we can find that for the same *c*, the estimation of λ_{EBi} and λ_{HBi} (i=1, 2, 3) are very close. Moreover, we can find that for the different *c*, λ_{EBi} and λ_{HBi} (i=1, 2, 3) are all robust.

TABLE II
THE ESTIMATE OF λ_{EBi} AND λ_{HBi}

c	λ_{EB1}	λ_{EB2}	λ_{EB3}	λ_{HB1}	λ_{HB2}	λ_{HB3}
0.1	1.0150	1.0155	1.0144	1.0140	1.0137	1.0141
0.2	1.0133	1.0144	1.0122	1.0134	1.0137	1.0132
0.3	1.0116	1.0133	1.0099	1.0123	1.0132	1.0117
0.4	1.0099	1.0122	1.0077	1.0110	1.0124	1.0099
0.5	1.0083	1.0111	1.0055	1.0099	1.0115	1.0081
0.6	1.0066	1.0100	1.0033	1.0081	1.0106	1.0062
0.7	1.0050	1.0089	1.0011	1.0066	1.0096	1.0043
0.8	1.0033	1.0077	0.9989	1.0052	1.0086	1.0024
0.9	1.0017	1.0066	0.9968	1.0037	1.0077	1.0005
1.0	1.0001	1.0056	0.9946	1.0023	1.0067	0.9986
range	0.0149	0.0099	0.0198	0.0117	0.007	0.0155

According to Table II, let $\lambda = 1$, we can get the absolute error of λ_{EBi} and λ_{HBi} are $\Delta_{EBi}=|\lambda_{EBi}-\lambda|$ and $\Delta_{HBi}=|\lambda_{HBi}-\lambda|$, respectively. Some numerical results are listed in Table III.

From Table III, we can find that for the same *c* (0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0), the results of Δ_{EBi} and Δ_{HBi} (i=1, 2, 3) are very close; moreover, for the different *c* (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7,0.8,0.9,1.0) , $\Delta_{EB1} \in [0.0001, 0.0150]$, $\Delta_{EB2} \in [0.0056, 0.0155]$, $\Delta_{EB3} \in [0.0011, 0.0144]$, $\Delta_{HB1} \in [0.0023, 0.0140]$, $\Delta_{HB2} \in [0.0067, 0.0137]$, $\Delta_{HB3} \in [0.0014, 0.0141]$.

TABLE III
THE RESULTS OF Δ_{EBi} AND Δ_{HBi}

c	λ_{EB1}	λ_{EB2}	λ_{EB3}	λ_{HB1}	λ_{HB2}	λ_{HB3}
0.1	0.0150	0.0155	0.0144	0.0140	0.0137	0.0141
0.2	0.0133	0.0144	0.0122	0.0134	0.0137	0.0132
0.3	0.0116	0.0133	0.0099	0.0123	0.0132	0.0117
0.4	0.0099	0.0122	0.0077	0.0110	0.0124	0.0099
0.5	0.0083	0.0111	0.0055	0.0099	0.0115	0.0081
0.6	0.0066	0.0100	0.0033	0.0081	0.0106	0.0062
0.7	0.0050	0.0089	0.0011	0.0066	0.0096	0.0043
0.8	0.0033	0.0077	0.0011	0.0052	0.0086	0.0024
0.9	0.0017	0.0066	0.0032	0.0037	0.0077	0.0005
1.0	0.0001	0.0056	0.0054	0.0023	0.0067	0.0014

From the example, we can find that: (1) Focus on the E-Bayesian estimations, under different prior distributions, we can find that the MSE_{EBi} and $Bias_{EBi}$ (i=1, 2, 3) are very close to each other. (2) Focus on the hierarchical Bayesian estimations, under different priors distribution, we can find that the MSE_{HBi} and $Bias_{HBi}$ are very close to each other. (3) Focus on the robust of different estimator, we can find that the E-Bayesian estimations and hierarchical Bayesian estimations are very robust. Therefore, the E-Bayesian estimations and hierarchical Bayesian estimations are very close to each other. Considering that the hierarchical Bayesian estimations include complicated integrals, these estimations are not obtained explicitly. For these reasons, the E-Bayesian method is a good choice instead of the hierarchical Bayesian method. From the application example, we find that the E-Bayesian estimation method is both efficient and easy to perform.

IX. CONCLUSION

In this paper, we study the E-Bayesian and hierarchical Bayesian estimations of the parameter under entropy loss function. The formulas of E-Bayesian estimation and hierarchical Bayesian estimation of the Poisson distribution are provided. Moreover, the relationship between hierarchical Bayesian estimation and E-Bayesian estimation is also discussed. Finally, a real numerical example is provided to show that E-Bayesian estimation is much simpler than hierarchical Bayesian estimation in practice.

REFERENCES

[1] L. Simar. Maximum likelihood estimation of a compound Poisson process. The Annals of Statistics, pp. 1200-1209, 1976.

[2] S. M. Sadooghi-Alvandi. Estimation of the parameter of a Poisson distribution using a LINEX loss function. Australian Journal of Statistics, vol.32, no.3, pp. 393-398, 1990.

[3] D. V. Lindley. Smith A.F., Bayes estimation for the linear model, Journal of the Royal Statistical Society-Series B, vol.34, pp. 1-41, 1972.

[4] M. Han. The structure of hierarchical prior distribution and its applications, Chinese Operations Research and Management Science, vol.6, no.3, pp. 31-40, 1997.

[5] M. Han. E-Bayesian estimation and hierarchical Bayesian estimation of failure rate. Applied Mathematical Modelling, vol.33, no.4, pp. 1915-1922, 2009.

[6] M. Han. E-Bayesian estimation of the reliability derived from Binomial distribution. Applied Mathematical Modelling, vol.35, no.5, pp. 2419-2424, 2011.

[7] J. Wang, D. Li, D. Chen. E-Bayesian Estimation and Hierarchical Bayesian Estimation of the System Reliability Parameter. Systems Engineering Procedia, vol.3, no.1, pp. 282-289, 2012.

[8] F. Yousefzadeh. E-Bayesian and hierarchical Bayesian estimations for the system reliability parameter based on asymmetric loss function. Communications in Statistics-Theory and Methods, vol.46, no.1, pp. 1-8, 2017.

[9] M. Han. The E-Bayesian and hierarchical Bayesian estimations of Pareto distribution parameter under different loss functions. Journal of Statistical Computation and Simulation, vol.87, no.3, pp. 577-593, 2017.

[10] R. Calabria, G. Pulcini. On the maximum likelihood and least-squares estimation in the inverse Weibull distributions, Stat Appl. vol.2, no.1, pp. 53-66, 1990.

[11] CP LI, and HB Hao. “Likelihood and Bayesian estimation in stress strength model from generalized exponential distribution containing outliers”, IAENG International Journal of Applied Mathematics, vol.46, no.2, pp. 155-159, 2016.