Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 13, No. 1, 2024

Empirical Bayes Likelihood for Exponential Distribution Family under Ranked Set Sampling and Dependent Data

Naiyi Li[1], Yongming Li[2], Juan Huang[1, *]

[1]School of Mathematics and Computer, Guangdong Ocean University, Zhanjiang 524088, China
[2]School of Mathematics and Computer Science, Shang Rao Normal University Shang Rao, 334001, China
[*] Corresponding Author: Juan Huang

Abstract: In this paper, we get the empirical Bayes likelihood test rules for Exponential distribution based on ranked set sampling. Its asymptotic optimality is obtained. Due to the advantages of combining empirical Bayesian methods with prior information of samples, the accuracy of statistical inference can be improved. Therefore, the fusion of empirical Bayesian likelihood methods can further enhance the estimation methods of statistical models.

Keywords: Ranked set sampling, empirical Bayes likelihood, Exponential distribution.

1. Introduction

Empirical Bayesian method is a method that uses historical samples to estimate prior distributions and determine Bayesian decision functions. This method is particularly useful when the prior distribution is unknown. Empirical Bayesian methods can usually be divided into two types: parametric empirical Bayesian methods and non parametric empirical Bayesian methods.In the parameter empirical Bayesian method, it is assumed that the prior distribution belongs to a known parameter family. Then, based on independent and identically distributed historical samples, classical statistical methods are used to provide estimates of parameters, thereby obtaining estimates of prior distributions. Using its Bayesian solution as an empirical Bayesian decision function.In non parametric empirical Bayesian methods, the Bayesian solution of a decision problem can be expressed as the product of two known functions. At this point, independent samples that follow a certain distribution can be used to estimate these two functions, thereby estimating the Bayesian solution and using it as the empirical Bayesian decision function for the decision problem.

Overall, the empirical Bayesian method is a method that utilizes historical data to optimize decision-making. It combines prior information and sample information, obtains posterior information based on Bayesian formulas, and then infers unknown parameters based on the posterior information Ranked set sampling (RSS) was introduced by McIntyre to estimated pasture yields[1]. RSS has been used in many fields such as environment and sociology. Since Robbins proposed empirical Bayes (EB) approach, it has been developed [2-9].

In this paper, we obtain empirical Bayes likelihood test rule for Exponential distribution based on ranked set sampling.

Let X have a conditional density function for given θ

$$f\left(x\mid\theta\right)=m\left(x\right)h\left(\theta\right)\exp\left\{-\theta x\right\} \tag{1.1}$$

where θ is unknown parameter, $x\in\Omega=\left\{-\infty\le a_1<x<a_2\le\infty\right\}$ is the sample space and $\Theta=\left\{\theta>0\mid 0<h^{-1}\left(\theta\right)=\int_\Omega m\left(x\right)e^{-\theta x}dx<\infty\right\}$ is parameter space.

We discuss the following test problem:

$$H_0:\theta\le\theta_0\leftrightarrow H_1:\theta>\theta_0 \tag{1.2}$$

Where H0 is given constants.

We choose loss function $L_0\left(\theta,d_0\right)=a\dfrac{\left(\theta-\theta_0\right)}{\theta}I\left(\theta>\theta_0\right)$, $L_1\left(\theta,d_1\right)=a\dfrac{\left(\theta-\theta_0\right)}{\theta}I\left(\theta\le\theta_0\right)$.

where $a>0,\ d=\{d_0,d_1\}$ is action space, d_0 *and* d_1 imply acceptance and rejection of h_0 respectively.

Suppose that the prior distribution $G(\theta)$ of parameter the θ is unknown.

We have random decision function

$$\delta(x)=P(\text{accept } H_0 \mid X=x). \tag{1.3}$$

Then, the risk function of δ(x) is shown by

$$\begin{aligned}R\left(\delta\left(x\right),G\left(\theta\right)\right)&=\int_\Theta\int_\Omega\Big[L_0\left(\theta,d_0\right)f\left(x\mid\theta\right)\delta\left(x\right)+\\&\qquad L_1\left(\theta,d_1\right)f\left(x\mid\theta\right)\left(1-\delta\left(x\right)\right)\Big]dxdG\left(\theta\right)\\&=a\int_\Omega\beta\left(x\right)\delta\left(x\right)dx+C_G\end{aligned}$$

where

$$C_G=\int_\Omega L_1\left(\theta,d_1\right)dG\left(\theta\right),$$
$$\beta\left(x\right)=\int_\Theta\frac{\left(\theta-\theta_0\right)}{\theta}f\left(x\mid\theta\right)dG\left(\theta\right). \tag{1.4}$$

The marginal density function of X is shown by

$$f_G(x)=\int_\Theta f(x\mid\theta)\,dG(\theta)$$
$$=\int_\Theta m(x)h(\theta)e^{-\theta x}dG(\theta)=m(x)p(x)$$

where $p(x)=\int_\Theta h(\theta)e^{-\theta x}dG(\theta)$.

Since $\frac{1}{\theta}h(\theta)e^{-\theta z}=\int_z^\infty h(\theta)e^{-\theta y}dy$,we have

$$\int_\Theta \frac{1}{\theta}f(x\mid\theta)\,dG(\theta)=m(x)\int_\Theta\int_x^\infty h(\theta)e^{-\theta y}dy\,dG(\theta)$$
$$=m(x)\int_x^\infty p(y)\,dy=m(x)\phi_G(x)$$

$$\phi_G(x)=\int_x^\infty p(y)\,dy=\int_\Theta \frac{1}{\theta}h(\theta)e^{-\theta x}dG(\theta), \quad (1.5)$$

By (1.5), we have $\beta_n(x)=f_n(x)-\theta_0 m(x)\phi_n(x)$ where

$$\phi_G(x)=\int_x^\infty p(y)\,dy=\int_x^\infty \frac{f_G(y)}{m(y)}dy=E\left(\frac{I_{\{X>x\}}}{m(x)}\right).$$

Using (1.5), Bayes test function is obtained as follows

$$\delta_G(x)=\begin{cases}1,\beta(x)\le 0,\\ 0,\beta(x)>0.\end{cases}$$

Further, we obtian the minimum Bayes risk as follows

$$R(G)=\inf_\delta R(\delta,G)=R(\delta_G,G)$$
$$=a\int_\Omega \beta(x)\delta_G(x)\,dx+C_G. \quad (1.6)$$

From above that $\delta(x)=\delta_G(x)$ and R(G) can be obtained when the prior distribution of $G(\theta)$ is given. If not, we use the EB method. The rest of this paper is organized as follows. Section 2 presents an EB test under ranked set sampling. In section 3, we obtain the optimal rate of convergence of the empirical Bayes likelihood test under ranked set sampling.

2. Construction of Empirical Bayes Likelihood Test under Ranked Set Sampling and dependent data

Supposed that $X_{(1)1}, X_{(1)2}, \cdots, X_{(1)m}, X_{(2)1}, X_{(2)2}\cdots, X_{(2)m}, \cdots, X_{(k)1}, X_{(k)2}\cdots, X_{(k)m}$ be a balanced ranked set sample from population which has the common marginal density function $f_{G(x)}$. We assume perfect ranking. Denote that $X_{(1)1}, X_{(1)m}, X_{(2)1}, X_{(2)m}, X_{(k)1}, X_{(k)m}$ are ranked set historical samples, and X is present sample. Assume $f(x)\in C_{s,\alpha}$, $x\in R_1$, where $C_{s,\alpha}=\{g(x)|g(x)$ is a probability density function. The s−he order derivative Supposed that K(x) be a Borel measurable bounded function vanishing off (0,1) such that (A1):

$$\frac{1}{t!}\int_0^1 v^t K(v)\,dv=\begin{cases}1,t=0\\ 0,t=1,\cdots,s-1\end{cases} \quad (2.1)$$

empirical likelihood is established by

$$R(f(x))=\sup\left\{\prod_{i=1}^n np_i, p_i\ge 0, \sum_{i=1}^n p_i=1, \sum_{i=1}^n p_i\Delta_i=0\right\}$$

empirical likelihood ratio is established by

$$l(f(x))=-2\log R(f(x))=2\sum_{i=1}^n \log(1+s\Delta_i).$$

where $s\in R$, s is determined by

$$\Pi(s)=\frac{1}{n}\sum_{i=1}^n\frac{\Delta_i}{1+s\Delta_i}=0.$$

Kernel estimator of f(x) is defined by

$$\Delta_i=K(x-X_{(i)1}/h_n)-f(x)$$

where h_n is a positive and smoothing bandwidth, and $\lim_{n\to\infty}h_n=0$, $f_n(x)=\arg\max R(f(x))$.

Thus, the estimator of $\beta(x)$ is shown by

$$\beta_n(x)=f_n(x)-\theta_0 m(x)\phi_n(x) \quad (2.2)$$

And, the EB test function is defined as follows

$$\delta_n(x)=\begin{cases}1,\beta_n(x)\le 0\\ 0,\beta_n(x)>0\end{cases} \quad (2.3)$$

Let E stand for mathematical expectation with respect to the joint distribution of $X_{(1)1}, X_{(1)2}, \cdots, X_{(1)m}, X_{(2)1}, X_{(2)2}, \cdots, X_{(2)m}, \cdots, X_{(k)1}, X_{(k)2}\cdots, X_{(k)m}$. Then, the overall Bayes risk of δ_n is shown by

$$R(\delta_n(x),G)=a\int_\Omega \beta(x)E_n[\delta_n(x)]\,dx+C_G$$

If $R(\delta_n,G)-R(\delta_G,G)=O(n^{-q})$, where $q>0$, $O(n^{-q})$ is asymptotic optimalit convergence rates of empirical Bayes likelihood test function { $\delta_n(x)$ }. Before proving the theorems, we need the following lemmas. Supposed that c,c1 be different constants in different cases even in the same expression.

Lemma. $R(\delta_G,G)$ and $R(\delta_n,G)$ are defined by above, then $0\le R(\delta_n,G)-R(\delta_G,G)\le c\int_\Omega|\beta(x)|P(|\beta_n(x)-\beta(x)|\ge|\beta(x)|)dx$.

3. Conclusion

Theorem Assume (C1) and the following regularity conditions hold. where $h_n = n^{-1/(2+s)}$,

(1) $\int_{\Omega} \left|\beta(x)\right|^{1-\lambda} m^{\lambda}(x)\, dx < \infty$,

(2) $\int_{\Omega} \left|\beta(x)\right|^{1-\lambda} dx < \infty$.

Then, we get $R(\delta_n, G) - R(\delta_G, G) = O\left(n^{-\lambda(s+1)/2(s+2)}\right)$

Proof. Using Markov's inequality, we get

$$\begin{aligned} 0 &\le R(\delta_n, G) - R(\delta_G, G) \\ &\le a\int_{\Omega} \left|\beta(x)\right|^{1-\lambda} E\left|\beta_n(x) - \beta(x)\right|^{\lambda} dx \\ &\le c\int_{\Omega} \left|\beta(x)\right|^{1-\lambda} E\left|f_n(x) - f_G(x)\right|^{\lambda} dx + \\ &\quad c\int_{\Omega} \left|\beta(x)\right|^{1-\lambda} m^{\lambda}(x) E\left|\phi_n(x) - \phi_G(x)\right|^{\lambda} dx \\ &\equiv \mathrm{M}_n + \Gamma_n \end{aligned} \tag{3.1}$$

Applying Lemma and the conditions (1)-(2) in the Theorem, we obtain

$$\mathrm{M}_n \le c \cdot n^{-\lambda(s+1)/2(s+2)} \int_{\Omega} \left|\beta(x)\right|^{1-\lambda} dx \le c \cdot n^{-\lambda(s+1)/2(s+2)} \tag{3.2}$$

$$\Gamma_n \le c \cdot n^{-\lambda/2} \int_{\Omega} \left|\beta(x)\right|^{1-\lambda} m^{\lambda}(x)\, dx \le c \cdot n^{-\lambda/2} \tag{3.3}$$

Substituting (3.2)-(3.3) into (3.1),

Remark 3.1. When $\lambda \to 1$, $s \to \infty$, $O\left(n^{-\lambda(s+1)/2(s+2)}\right)$ nears $O\left(n^{-1/2}\right)$.

Acknowledgements

This paper was financially supported by Natural Science Foundation of China (12161075), Natural Science Foundation of Guangdong Province (2022A1515010978, 2024A1515011258) Natural Science Foundation of Jiangxi Province (20122ACB201006) and Natural Science Foundation of Guangdong Ocean University (R17083, C17201, P16091).

References

[1] Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. Biometrika, 75(2):237-249.

[2] Owen, A. B. (1990). Empirical likelihood confidence regions. Annals of Statistics, 18(1): 90-120.

[3] Chen, S. X. (2000). Empirical likelihood confidence intervals for linear regression coefficients. Journal of Multivariate Analysis, 72(1):19-32.

[4] Kitamura, Y. (2001). Empirical likelihood methods with weakly dependent processes. Annals of Statistics, 29(1):116-139.

[5] Qin, J. and Lawless, J. (1994). Empirical likelihood and general estimating equations. Annals of Statistics, 22(1):300-325.

[6] McIntyre G A. (1952). A Method for Unbiased Selective Sampling, Using Ranked Sets (Australian Journal of Agricultural Research) , 3: 385–390.

[7] Chen Z H, Bai Z D, Sinha B K. Ranked Set Sampling: Theory and Applications (Springer 2003).

[8] Robbins, H. (1955). An empirical Bayes approach to statistics, Proc. Third Berkeley Symp. Math. Statist. Prob, Vol. 1: 157--163.

[9] Karunamuni, R. J, Li, J, Wu, J. (2010). Robust empirical Bayes tests for continuous distributions, Journal of Statistical Planning and Inference, 140 (1): 268-282.

[10] Xinyi Xu, Dunke Zhou. (2011). Empirical Bayes predictive densities for high-dimensional normal models, Journal of Multivariate Analysis, 102(10): 1417-1428.

[11] Johns M VJr, Van Ryzin J R. (1972). Convergence Rates in Empirical Bayes Two-action Problems 2: Continuous Case. Ann. Math. Statist. Vol. 43: 934-937.

[12] Lichun Wang, Radhey S. Singh. (2014). Linear Bayes estimator for the two-parameter exponential family under type II censoring. Computational Statistics & Data Analysis, Vol.71: 633-642.

[13] Rekab Ki. (1994). Modified shrinkage estimates in the exponential family case. Stochastic Analysis and Applications, Vol. 4: 493-498.

[14] Shang S, Li J. (2005). One empirical Bayes procedures for selecting good populations in a positive exponential family. Journal of Statistical Planning and Inference, Vol.129: 3-18.

[15] Naiyi Li, Yuan Li, Yongming Li, Yang Liu. (2015). Empirical Bayes Inference for the Parameter of Power Distribution Based on Ranked Set Sampling. Discrete Dynamics in Nature and Society: 1-5.