Estimation of Hyper-Parameters Based on Logit-Normal Model

Zehai ZHAO
Guizhou University, Guiyang 550025 P. R. China; 495359612@qq.com

Mingzhong JIN
Guizhou Minzu University, Guiyang 550025 P. R. China; 13885196899@163.com

Minqing GONG, Siwen YANG
Guizhou University, Guiyang 550025 P. R. China
minquing@163.com, 352333464@qq.com

ABSTRACT: In this paper, we deduce a range of the hyper-parameters, and use Gauss-Hermite quadrature introduced in reference [1] to approximate the moment equation, estimating the moment estimators by enumeration method. Simulation experiment carried out to study the accuracy of estimates by Monte Carlo integration. Simulation experiment shows that estimating parameters from the moment equation of logit-normal model by enumeration method can get very accurate estimates, and verified the EB estimator is robust. We also use the method to estimate poverty incidences in Spain.
KEYWORD: SAE; Logit-normal model; EB; Enumeration method; Poverty incidence

1 INTRODUCTION

The commonly used methods for SAE includes synthetic estimation, composite estimation, empirical best linear unbiased prediction (EBLUP), empirical Bayes (EB) method and hierarchical Bayes (HB) method. Empirical Bayes method and hierarchical Bayes methods are applicable more generally in the sense of handling models for binary and count data……[1]. Hyper-parameters of EB models can usually be estimated by method of the maximum likelihood (ML) or the moment estimation, but ML method is computationally complex. Several models for EB estimation was given in reference[1], including logit-normal model. The moment equation of paramaters can be approximated by using Gauss-Hermite quadrature first, and then its roots be obtained by using Newton-Raphson iterative algorithm[1]. We proposed a feasible approach in this paper to calculate the hyper-parameters of logit-normal model.

2 LOGIT-NORMAL MODEL[1]

Consider a two stage model on the sample observations y_{ij}, that is

$y_{ij} = 0$ or 1, $j = 1,\cdots,n_i$; $i = 1,\cdots,m$.

In the first stage, we assume that the proportion of *i*-th area is

$p_i = \sum_j y_{ij} / \sum_j 1$,

and its sample total is

$y_i = \sum_{j\in S_i} y_{ij}$,

where S_i denotes the sample set of *i*-th area, n_i is the sample size,

$y_i \sim B(n_i, p_i)$, that is,

$$f(y_i \mid p_i) = \binom{n_i}{y_i} p_i^{y_i} (1-p_i)^{n_i-y_i} \tag{1}$$

In the second stage, assume that p_i satisfies the logit-normal model:

$$\operatorname{logit}(p_i) = \log[p_i/(1-p_i)] \sim N(\mu, \sigma^2) \tag{2}$$

According to (2), the p_i can be written as:

$$p_i(a(z)) = \exp(a(z))/(1+\exp(a(z))),$$

$$z \sim N(0,1)\ a(z) = \mu + \sigma z \tag{3}$$

According to (1) and (3), the Bayes estimator is:

$$\hat{p}_i^{\mathrm{B}} = \frac{E[h_1(a(z))\exp\{h_2(y_i, a(z))\}]}{E[\exp\{h_2(y_i, a(z))\}]} \tag{4}$$

where $\varphi(z)$ is the density of normal $N(0,1)$ distribution, and

$$h_1(a) = \exp(a)/(1+\exp(a)),$$

$$h_2(y_i, a) = ay_i - n_i \ln(1+\exp(a)) \tag{5}$$

Formula (4) can be evaluated by simulating samples from $N(0,1)$ or using numerical integration.

3 ESTIMATION OF PARAMETERS μ AND σ

3.1 *Moment equation for μ and σ [1].*

Equate the weighted sample mean and the weighted sample variance to $p_1,\cdots,p_m$, which leads to:

$$E(p)=\sum_i \frac{n_i}{n_T}\hat{p}_i=\hat{p},$$

$$S_p^2=\sum_i \frac{n_i}{n_T}(\hat{p}_i-\hat{p})^2,$$

$$n_T=n_1+\cdots+n_m \tag{6}$$

The expectation of p^2 is obtained from (6), that is

$$E(p^2)=[E(p)]^2+\mathrm{var}(p)=\hat{p}^2+S_p^2 \tag{7}$$

According to (3), (5), (6) and (7), the moment estimators, $\hat{\mu}$ and $\hat{\sigma}$, are given by

$$E[h_1(\mu+\sigma z)]=\hat{p},$$

$$E[h_1^2(\mu+\sigma z)]=\hat{p}^2+S_p^2 \tag{8}$$

Let $A=1-\hat{p}$, $B=A^2+S_p^2$,

then equation (8) is equivalent to

$$\int_{-\infty}^{+\infty}\frac{1}{1+\exp(\mu+\sigma x)}\varphi(x)\mathrm{d}x=A,$$

$$\text{and}\quad \int_{-\infty}^{+\infty}\frac{1}{(1+\exp(\mu+\sigma x))^2}\varphi(x)\mathrm{d}x=B \tag{9}$$

where $\varphi(x)$ is the density of the normal $N(0,1)$ distribution.

Since the $\hat{p}$ usually satisfies $\hat{p}<1/2$, that is, $A>1/2$, we assume that $(\hat{\mu},\hat{\sigma})$ is the solution of equation (9), then we have the following inequality:

$$\ln(2+B-3A)-\hat{\sigma}^2/2<\hat{\mu}<\ln(1/A-1) \tag{10}$$

3.2 *Enumeration method to obtain the estimates of μ and σ*

Assume that

$a<\hat{\mu}<b$, $c<\hat{\sigma}<d$, let

$\mu_i=a+(b-a)(i/N_1)$, $i=1,\cdots,N_1$,

$\sigma_j=c+(d-c)(j/N_2)$, $j=1,\cdots,N_2$,

$\boldsymbol{\alpha}=(\mu_1,\cdots,\mu_{N_1})^T$,

$\boldsymbol{\beta}=(\sigma_1,\cdots,\sigma_{N_2})^T$,

then the *i*-th row and *j*-th column of matrix

$\boldsymbol{\alpha}1_{N_2}^T$ and $1_{N_1}\boldsymbol{\beta}^T$

is corresponding to (μ_i,σ_j).

For all i, j, and positive number d_1 and d_2 which are close to 0, we use the Gauss-Hermite quadrature [1] to approximate the left of equation (9), and pick off μ_i and σ_j which satisfy the following inequality:

$$\left|\int_{-\infty}^{+\infty}\frac{1}{1+\exp(\mu+\sigma x)}\varphi(x)\mathrm{d}x-A\right|<d_1,$$

$$\left|\int_{-\infty}^{+\infty}\frac{1}{(1+\exp(\mu+\sigma x))^2}\varphi(x)\mathrm{d}x-B\right|<d_2 \tag{11}$$

We use the mean of μ and σ or the intersection of two “straight lines” respectively to be the estimates. Substitute the moment estimators $\hat{\mu}$ and $\hat{\sigma}$ into (3), and use the Gauss-Hermite quadrature or MC method to get a EB estimator of p_i.

4 SIMULATION EXPERIMENT AND RESULTS ANALYSIS

Assume that there are 50 small areas with 50 samples, that is, $m=50$, $n_i=50$.

Let the sample total for the *i*-th area be y_i, and the probability of $y_{ij}=1$ be 0.1, namely, $y_i\sim B(50,0.1)$.

Let $n=20$ for the Gauss-Hermite quadrature. The direct estimator of p_i is

$\hat{p}_i=y_i/n_i$.

We do the simulation experiment by MATLAB.

1) Generate 50 binomially distributed random numbers, and compute A and B as follows:

y=(7 3 7 8 6 5 5 4 1 6 3 4 3 4 9 3 3 1 5 6 4 3 1 5 3 4 4 3 3 4 6 8 3 7 6 4 1 6 5 3 6 4 7 5 10 5 5 5 6 4); A =0.9068; B =0.823848.

2) Assume that $\sigma\in(\sigma_0,\sigma_1)$ and select the initial value:

$\sigma_0=0$, $\sigma_1=1$,

the interval of μ is determined by (10). The intervals of μ and σ are divided into 1000 equal sections, let

$$d_1=d_2=10^{-5},$$

and use the Gauss-Hermite quadrature to

approximate the left of equation (10), the image for (11) is shown in Figure 1.

Reduce the range of μ and σ:

$\sigma_0 = 0.45$, $\sigma_1 = 0.55$, and let $d_1 = d_2 = 5\times10^{-6}$,

the image of (11) is shown in Figure 2.

Re-adjust the range of μ and σ again:

$\sigma_0 = 0.45$, $\sigma_1 = 0.47$.

3) Let $\sigma_0 = 0.45$, $\sigma_1 = 0.47$,

$\mu_0 =$ -2.37 , $\mu_1 =$ -2.35 $d_1 = d_2 = 5\times10^{-7}$,

intervals of μ and σ are divided into 1000 equal sections, and select the nodes which satisfy (11), we get the nodes:

x=492, 492, 491; y=561, 562, 565,

the estimates of μ and σ are calculated as:

$\hat{\mu} =$-2.36017, $\hat{\sigma} = 0.46123$.

We substitute the moment estimators

$\hat{\mu} =$-2.36017, $\hat{\sigma} = 0.46123$

into the left of following (12) and (13) by generating 100000 normally distributed random numbers

$x_1, \cdots, x_{100000}$,

we get 6 groups of values as shown in Table 1.

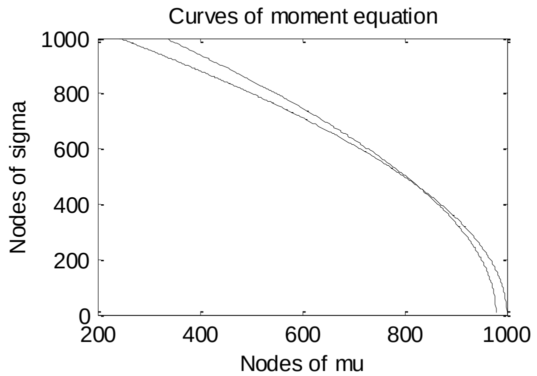

Figure 1. Curves of parameters $\sigma_0 = 0$ $\sigma_1 = 1$.

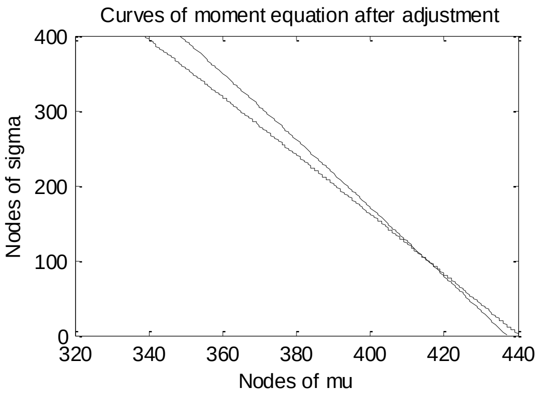

Figure 2. Curves of parameters $\sigma_0 = 0.45$ $\sigma_1 = 0.55$.

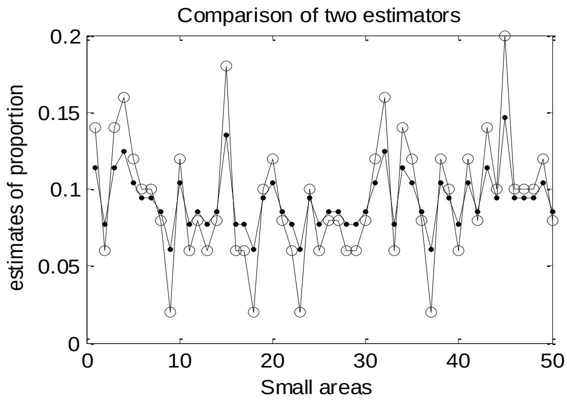

Figure 3. Comparison of two estimators.

$$\frac{1}{200000}\left(\sum_{i=1}^{100000}\frac{1}{1+\exp(\hat{\mu}+\hat{\sigma}x_i)}+\frac{1}{1+\exp(\hat{\mu}-\hat{\sigma}x_i)}\right)-A \qquad (12)$$

$$\frac{1}{200000}\left(\sum_{i=1}^{100000}\frac{1}{(1+\exp(\hat{\mu}+\hat{\sigma}x_i))^2}+\frac{1}{(1+\exp(\hat{\mu}-\hat{\sigma}x_i))^2}\right)-B \qquad (13)$$

We substitute the moment estimators

$\hat{\mu} =$-2.36017, $\hat{\sigma} = 0.46123$

into (3), and use the Gauss-Hermite quadrature to get the EB estimator of p_i.The scatter diagram of $\hat{p}_i$ and $\hat{p}_{i;PEB}$ is shown in Figure 3('o' corresponds to direct estimates, '·' corresponds to EB estimates).

Table 1. Values of formula (12) and (13). unit:10^{-5}

Value of (12)	3.10	-0.12	-1.23	-0.08	-1.25	2.90
Value of (13)	4.71	-0.09	-1.84	-0.29	-1.88	4.59

5 EXAMPLE

The EB method based on logit-normal is applied to compute poverty incidences in Spanish provinces. The data of n_d and $\hat{p}^{\varpi}$ come from the European Survey on Income and Living Conditions for the year 2006[2], we let the y_d is close to $n_d \cdot \hat{p}^{\varpi}$, and get y_d to calculate the $\hat{p}^{\mathrm{EB}}$, which is shown in Table 2.

Table 2. Direct and EB estimates of poverty incidence.

Province	n_d	y_d	$\hat{p}^{\varpi}(\%)$	$\hat{p}^{EB}(\%)$
Soria	17	10	60.41	45.34
Tarragona	129	16	12.46	12.66
Cordoba	230	71	30.66	30.38
Badajoz	427	173	36.58	36.31
Barcelona	1482	161	10.82	10.91

The estimates of μ and σ are calculated as:

$\hat{\mu}=$-1.66964; $\hat{\sigma}=$ 0.80218.

Substitute the moment estimators $\hat{\mu}$ and $\hat{\sigma}$ into (3), and use MC integration as in part 4 to get the EB estimates of p_i as shown in Table 2. Values of $\hat{p}^{EB}$ in Table 2 show that the estimates is adjusted to a reasonable "line" by using EB method, especially for the small sample size. And the EB estimators are closer to direct estimators with the sample size increasing.

6 SUMMARY

By Monte Carlo method we have: It would be an efficient numerical method to use Gauss-Hermite quadrature to approximate the moment equation of logit-normal model .It would be easy to find a short interval of μ and σ by observing the image of parameters and estimate the hyper-parameters by using enumeration method; The EB estimators based on logit-nomal model is robust. 4). Estimation of proportion for rare events may be not feasible because the curves of moment equation is almost coincident. In this paper inspired by [2] we use different EB method from [2] to calculate the real data example in [2] and our way is shown feasible also.

7 ACKNOWLEDGMENTS

National Statistics Bureau of China (major program 2012LZ054). Guizhou Provincial Governor Funding (QianKeJiaoBan[2010]-04). Guiyang Bureau of Sci. and Tech. (ZhuKeHeTong[2011101]1-32). Guizhou Dept. of Sci. and Tech. (QianKeHeJiZi[2009]4002).

REFERENCES

[1] J.N.K.RAO. Small Area Estimation. American: Wiley, 2003: 197-201.

[2] Isable Molina and J.N.K.RAO. Small Area Estimation of Poverty Indicators. The Canadian Journal of Statistics, 2010(38):369-385.

[3] Yongjin Jin. Sampling Survey: Theory and Application. China: Higher Education Press.

[4] Christian P.Robert. Monte Carlo Methods. Springer Tests in Statistics.

[5] Danny Pfeffermann. New Important Developments in Small Area Estimation. Statistical Science, 2013, Vol.28 (No.1).

[6] L.Wei, W. Zhang. Bayesian Analysis. Press of University of Science and Technology of China