第43卷第2期 中国科学技术大学学报 Vol. 43, No. 2
2013年2月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Feb. 2013

文章编号:0253-2778(2013)02-0162-07

正态总体误差方差的经验 Bayes 估计及其优良性

杨奉豪,冯 珉

(中国科学技术大学统计与金融系,安徽合肥 230026)

摘要:对正态总体误差方差在共轭先验分布和加权平方损失下导出了其 Bayes 估计,构造了其参数型经验 Bayes(PEB)估计,研究了其在均方误差(MSE)准则下相对于一致最小方差无偏估计(UMVUE)的优良性. 当先验分布中的超参数完全未知时,通过数值模拟比较了 PEB 估计和 UMVUE 的均方误差,获得了 PEB 估计的优良性.

关键词:误差方差;正态分布;PEB 估计;MSE 准则

中图分类号:O212.1 **文献标识码**:A doi:10.3969/j.issn.0253-2778.2013.02.012

AMS Subject Classification (2000): 62C12

引用格式: Yang Fenghao, Feng Min. The empirical Bayes estimation and its superiority for error variance in normal distribution[J]. Journal of University of Science and Technology of China, 2013,43(2):162-168.
杨奉豪,冯珉. 正态总体误差方差的经验 Bayes 估计及其优良性[J]. 中国科学技术大学学报,2013,43(2):162-168.

The empirical Bayes estimation and its superiority for error variance in normal distribution

YANG Fenghao, FENG Min

(*Department of Statistics and Finance, University of Science and Technology of China, Hefei 230026, China*)

Abstract: Under the conjugate prior distribution of the error variance in normal distribution and the weighted squared error loss function, the Bayes estimator was derived and the parametric empirical Bayes (PEB) estimator was constructed for the error variance. The superiority of the PEB estimation over the uniformly minimum variance unbiased estimation (UMVUE) in terms of the mean-square error (MSE) criterion was studied. In the case where the hyper-parameters of the prior distribution are completely unknown, the superiority of the PEB estimation over the UMVUE under the MSE criterion was investigated with a simulation study.

Key words: error variance; normal distribution; PEB estimation; MSE criterion

0 引言

在实际问题中,有很多的试验数据服从正态分布. 中心极限定理保证,很多分布虽然不是严格服从正态分布,但是渐近服从正态分布. 故对正态分布参数估计的研究具有重要意义. 文献中对正态总体均值的研究已经比较多了. 本文主要研究正态总体的方差. 当对误差方差除了有样本信息之外还具有先

收稿日期:2012-08-13;**修回日期**:2013-01-04

基金项目:国家重点基础研究发展(973)计划(2007CB814901),国家自然科学基金(11171179)资助.

作者简介:杨奉豪(通讯作者),男,1988 年生,研究生. 研究方向:概率论与数理统计. E-mail: yfh@mail.ustc.edu.cn

验信息时,引入 Bayes 方法[1-2] 可以提高估计的效果. 在许多实际问题中,我们对参数的先验知识有一定的积累,然而这种先验知识往往还不足以形成先验分布. 此时,如果人为地指定一个先验分布但却与实际情形偏离较大,则 Bayes 估计的性质会较差. 经验 Bayes 方法就是针对这一问题而提出的,它的实质是利用历史样本对先验分布或其数字特征作出估计.

按 Morris[3],统计模型中参数的经验 Bayes (Empirical Bayes,EB)方法可分为两类:非参数的经验 Bayes(NPEB)和参数型的经验 Bayes(PEB)方法. NPEB 方法通常假定先验分布的形式未知,但先验分布的某些矩存在,在一定损失函数下,求得的有关参数的 Bayes 估计常常表达为概率密度函数及其偏导(或分布函数)的函数,利用历史样本采用非参数方法对概率密度函数及其偏导(或分布函数)作出估计,从而获得 NPEB 估计. 对 NPEB 估计,通常研究其大样本性质,如渐近正态性或收敛速度等[4-6].

PEB 方法通常假定先验分布的形式已知,但含有未知的超参数和多余参数,利用历史样本对超参数和多余参数作出估计,从而获得 PEB 估计.

对 PEB 估计通常研究其小样本性质,如 MSE 准则下的优良性、稳健性和可容性等[7-9].

本文将研究正态分布误差方差的 PEB 估计及其优良性问题,而文献中对这一问题研究较少.

关于正态分布,我们还需要以下准备知识:

设随机变量(r. v.)X 的概率分布为正态分布,其密度函数为

$$f(x \mid \theta,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2\sigma^2}(x-\theta)^2\right\} \tag{1}$$

我们取 σ^2 的先验分布为逆 Gamma 分布 $\Gamma^{-1}(r/2,\lambda/2)$,其密度函数为

$$\pi(\sigma^2) = \frac{(\lambda/2)^{r/2}}{\Gamma(r/2)}(\sigma^2)^{-\left(\frac{r}{2}+1\right)}\exp\left\{-\frac{\lambda}{2\sigma^2}\right\} \propto (\sigma^2)^{-\left(\frac{r}{2}+1\right)}\exp\left\{-\frac{\lambda}{2\sigma^2}\right\} \tag{2}$$

式中,$\sigma^2>0$,而 r,λ 称为先验分布的超参数;“$\propto$”表示“正比于”. 设 $X_1,X_2,\cdots,X_m$ 为从总体(1)中抽取的 i. i. d(独立同分布)样本,记 $\boldsymbol{X}=(X_1,\cdots,X_m)$,则 σ^2 的似然函数为

$$l(\sigma^2,\boldsymbol{x}) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^m\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^m(x_i-\theta)^2\right\} \propto (\sigma^2)^{-m/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^m(x_i-\theta)^2\right\} \tag{3}$$

由式(2)和式(3)可知 σ^2 的后验密度为:

$$\pi(\sigma^2 \mid \boldsymbol{x}) \propto l(\sigma^2 \mid \boldsymbol{x})\pi(\sigma^2) \propto (\sigma^2)^{-\left(\frac{m+r}{2}+1\right)}\exp\left\{-\frac{\lambda+A}{2\sigma^2}\right\} \tag{4}$$

此处 $A=\sum_{i=1}^m(x_i-\theta)^2$. 可见后验分布 $\pi(\sigma^2 \mid \boldsymbol{x})$ 服从逆 Gamma 分布 $\Gamma^{-1}((m+r)/2,(\lambda+A)/2)$. 添加正则化常数,得后验密度为

$$\pi(\sigma^2 \mid \boldsymbol{x}) = \frac{[(\lambda+A)/2]^{(m+r)/2}}{\Gamma((m+r)/2)}(\sigma^2)^{-\left(\frac{m+r}{2}+1\right)}\exp\left\{-\frac{\lambda+A}{2\sigma^2}\right\} \tag{5}$$

令损失函数为

$$\left.\begin{aligned} &L(d,\sigma^2) = \left(\frac{d-\sigma^2}{\sigma^2}\right)^2 = \omega(\sigma^2)(d-\sigma^2)^2 \\ &\omega(\sigma^2) = \sigma^{-4}\end{aligned}\right\} \tag{6}$$

加权平方损失下 σ^2 的 Bayes 估计为

$$\hat{\sigma}^2_{\mathrm{B}} = \frac{E(\omega(\sigma^2)\sigma^2 \mid \boldsymbol{x})}{E(\omega(\sigma^2) \mid \boldsymbol{x})} = \frac{E(\sigma^{-2} \mid \boldsymbol{x})}{E(\sigma^{-4} \mid \boldsymbol{x})} = \frac{\sum_{i=1}^m(X_i-\overline{X})^2+m(\overline{X}-\theta)^2+\lambda}{m+r+2} \tag{7}$$

而经典统计方法中 σ^2 的一致最小方差无偏估计(UMVUE)是[10]

$$\hat{\sigma}^2 = S^2 = \frac{1}{m-1}\sum_{i=1}^m(X_i-\overline{X})^2 \tag{8}$$

本文的节 1 中,我们将对 r 已知而 λ,θ 未知的情形,构造 σ^2 的 PEB 估计. 节 2 中,我们在均方误差准则下研究 σ^2 的 PEB 估计相对于 UMVUE 的优良性. 节 3 中将对 r 和 λ,θ 皆未知的情形,我们重新构造 σ^2 的 PEB 估计,利用模拟方法获得 σ^2 的 PEB 估计相对于 UMVUE 的优良性.

1 当部分超参数未知时 PEB 估计的构造

本节将在 r 已知而 λ 和多余参数 θ 未知的情形下,构造 λ 的 PEB 估计.

由于 λ 和 θ 未知,由式(7)给出的 Bayes 估计无实用价值,我们要引入经验 Bayes 方法,利用历史样本对 θ,λ 作出估计从而获得 PEB 估计. 下面构造 σ^2 的 PEB 估计. 在 EB 估计问题的结构中我们常假定

$$(\mathbf{X}^{(1)},\sigma_1^2),(\mathbf{X}^{(2)},\sigma_2^2),\cdots,(\mathbf{X}^{(n)},\sigma_n^2),(\mathbf{X}^{(n+1)},\sigma_{n+1}^2)=(\mathbf{X},\sigma^2)$$

为相互独立的随机变量的对子；其中 $\sigma_1^2,\cdots,\sigma_n^2$ 和 $\sigma_{n+1}^2=\sigma^2$ 是不可观察的，但假定它们具有共同的先验分布 $\Gamma^{-1}(r/2,\lambda/2)$，并记 $\tilde{\sigma}^2=(\sigma_1^2,\cdots,\sigma_n^2,\sigma^2)$；而

$$\mathbf{X}^{(i)}=(X_{i1},\cdots,X_{im}),i=1,2,\cdots,n,$$
$$\mathbf{X}^{(n+1)}=\mathbf{X}=(X_1,\cdots,X_m)$$

是可观察的，假定它们具有共同的边缘分布[2].

当 $\mathbf{X}^{(i)}=(X_{i1},\cdots,X_{im})$ 为从 $N(\theta,\sigma_i^2)$ 中抽取的简单样本时，$\overline{X}_i=\frac{1}{m}\sum_{k=1}^{m}X_{ik}$ 和 $S_i^2=\frac{1}{m-1}\sum_{k=1}^{m}(X_{ik}-\overline{X}_i)^2$ 为充分统计量[10]，$i=1,2,\cdots,n+1$. 因此 $(\overline{X}_i,S_i^2)$，$i=1,2,\cdots,n+1$ 是 i.i.d 的样本，其中 $(\overline{X}_{n+1},S_{n+1}^2)=(\overline{X},S^2)$. 可用它们代替 $\mathbf{X}^{(i)}=(X_{i1},\cdots,X_{im})$，$i=1,2,\cdots,n+1$. 通常称 $(\overline{X}_1,S_1^2),\cdots,(\overline{X}_n,S_n^2)$ 为历史样本，称 $(\overline{X},S^2)$ 为当前样本. θ 用历史样本估计，即

$$\hat{\theta}=\frac{1}{n}\sum_{i=1}^{n}\overline{X}_i=\overline{\overline{X}} \tag{9}$$

由于 $\left.\frac{(m-1)S^2}{\sigma^2}\right|\tilde{\sigma}^2\sim\chi_{m-1}^2$，

$$E(S^2)=E[E(S^2\mid\tilde{\sigma}^2)]=E[\sigma^2]=\int\sigma^2\cdot\pi(\sigma^2)\mathrm{d}\sigma^2=\frac{\lambda}{r-2} \tag{10}$$

类似有

$$E(S_1^2)=E(S_2^2)=\cdots=E(S_n^2)=\frac{\lambda}{r-2}.$$

用 $\overline{S}^2=\frac{1}{n}\sum_{i=1}^{n}S_i^2$ 代替 $E(S^2)$，解方程(10)得到 λ 的矩估计量为

$$\hat{\lambda}_n=(r-2)\overline{S}^2=\frac{r-2}{n}\sum_{i=1}^{n}S_i^2 \tag{11}$$

将式(7)中 θ 和 λ 用 $\hat{\theta}$ 和 $\hat{\lambda}_n$ 代入，得 σ^2 的 PEB 估计为

$$\hat{\sigma}_{\mathrm{EB}}^2=\frac{1}{m+r+2}[(m-1)S^2+m(\overline{X}-\hat{\theta})^2+\hat{\lambda}_n] \tag{12}$$

2 PEB 估计在 MSE 准则下的优良性

为研究在 MSE 准则下 σ^2 的 PEB 估计相对于 UMVUE 的优良性，首先给出下列定义：

定义 2.1 设 $\hat{\theta}$ 为 θ 的一个估计量，则称 $\mathrm{MSE}(\hat{\theta})=E(\hat{\theta}-\theta)^2$ 为 $\hat{\theta}$ 的均方误差(MSE). 设 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 为 θ 的两个不同的估计量，若 $\mathrm{MSE}(\hat{\theta}_2)-\mathrm{MSE}(\hat{\theta}_1)>0$，则称在 MSE 准则下 $\hat{\theta}_1$ 优于 $\hat{\theta}_2$.

首先，我们为方便计算给出以下引理：

引理 2.1 设 $\tilde{\sigma}^2=(\sigma_1^2,\cdots,\sigma_n^2,\sigma^2)$ 中各分量相互独立，且具有共同的分布 $\Gamma^{-1}(r/2,\lambda/2)$，则有：

① $E(\sigma_i^2)=E(\sigma^2)=\frac{\lambda}{r-2}$，$E(\sigma_i^4)=E(\sigma^4)=\frac{\lambda^2}{(r-2)(r-4)}$；

② $E\left(\sum_{i=1}^{n}\sigma_i^2\right)^2=\frac{\lambda^2[(r-4)n^2+2n]}{(r-2)^2(r-4)}$；

③ $$E\left[\sigma^2+\frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2\right]^2=\frac{\lambda^2[(r-2)n^3+2(r-4)n^2+(r-4)n+2]}{n^3(r-2)^2(r-4)}.$$

证明 ①易证；

②由 $\tilde{\sigma}^2=(\sigma_1^2,\cdots,\sigma_n^2,\sigma^2)$ 中分量的独立性以及①得

$$E\left(\sum_{i=1}^{n}\sigma_i^2\right)^2=E\left(\sum_{i=1}^{n}\sigma_i^4+2\sum_{i<j}\sigma_i^2\sigma_j^2\right)=\frac{\lambda^2[(r-4)n^2+2n]}{(r-2)^2(r-4)};$$

③由 $\tilde{\sigma}^2=(\sigma_1^2,\cdots,\sigma_n^2,\sigma^2)$ 中分量的独立性以及①,②得

$$E\left[\sigma^2+\frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2\right]^2=E(\sigma^4)+\frac{2}{n^2}E\left(\sigma^2\sum_{i=1}^{n}\sigma_i^2\right)+\frac{1}{n^4}E\left(\sum_{i=1}^{n}\sigma_i^2\right)^2=\frac{\lambda^2[(r-2)n^3+2(r-4)n^2+(r-4)n+2]}{n^3(r-2)^2(r-4)}.$$

引理 2.2 设 $S_i^2=\frac{1}{m-1}\sum_{k=1}^{m}(X_{ik}-\overline{X}_i)^2$，$i=1,2\cdots,n+1$，则有：

① $E(S_i^2)=E(S^2)=\frac{\lambda}{r-2}$；

② $E(S_i^4)=E(S^4)=\frac{(m+1)\lambda^2}{(m-1)(r-2)(r-4)}$；

③ $$E\left(\sum_{i=1}^{n}S_i^2\right)^2=\frac{n\lambda^2[n(m-1)(r-4)+2r+2m-6]}{(m-1)(r-2)^2(r-4)}.$$

证明 ①由式(10)易证；

②由 $\left.\frac{(m-1)S_i^2}{\sigma^2}\right|\tilde{\sigma}^2\sim\chi_{m-1}^2$ 及引理 2.1①，得

$$E(S_i^4)=E(S^4)=E[E(S^4\mid\tilde{\sigma}^2)]=$$
$$E\left[\frac{\sigma^4}{(m-1)^2}\mathrm{Var}(\chi^2_{m-1}\mid\tilde{\sigma}^2)+\frac{\sigma^4}{(m-1)^2}(E(\chi^2_{m-1}\mid\tilde{\sigma}^2))^2\right]=$$
$$\frac{(m+1)\lambda^2}{(m-1)(r-2)(r-4)};$$

③由 $S_1^2,\cdots,S_n^2$ 的独立性以及①,②得

$$E\left(\sum_{i=1}^n S_i^2\right)^2=E\left(\sum_{i=1}^n S_i^4+2\sum_{i<j}S_i^2S_j^2\right)=$$
$$nES^4+n(n-1)(E(S^2))^2=$$
$$\frac{n\lambda^2[n(m-1)(r-4)+2r+2m-6]}{(m-1)(r-2)^2(r-4)}.$$

在 MSE 准则下，σ^2 的 PEB 估计相对于 UMVUE 的优良性有下列结果：

定理 2.1　设 σ^2 的 PEB 估计 $\hat{\sigma}^2_{EB}$ 和 UMVUE $\hat{\sigma}^2$ 分别由式(12)和式(8)给出，当 $r\geqslant 9$，$n>\max\left\{1,\frac{m}{10.5}\right\}$，$m>1$ 时，有

$$\mathrm{MSE}(\hat{\sigma}^2)-\mathrm{MSE}(\hat{\sigma}^2_{EB})>0\tag{13}$$

即在 MSE 准则下 σ^2 的 PEB 估计优于 UMVUE.

证明　接下来，将不等式(13)的左边部分做变换，有

$$\mathrm{MSE}(\hat{\sigma}^2)-\mathrm{MSE}(\hat{\sigma}^2_{EB})=$$
$$\mathrm{MSE}(\hat{\sigma}^2)-E(\hat{\sigma}^2_{EB}-\hat{\sigma}^2+\hat{\sigma}^2-\sigma^2)^2=$$
$$\mathrm{MSE}(\hat{\sigma}^2)-[E(\hat{\sigma}^2_{EB}-\hat{\sigma}^2)^2+2E[(\hat{\sigma}^2_{EB}-\hat{\sigma}^2)(\hat{\sigma}^2-\sigma^2)]+E(\hat{\sigma}^2-\sigma^2)^2]=$$
$$2E[(\hat{\sigma}^2-\hat{\sigma}^2_{EB})(\hat{\sigma}^2-\sigma^2)]-E(\hat{\sigma}^2-\hat{\sigma}^2_{EB})^2+\mathrm{MSE}(\hat{\sigma}^2)-E(\hat{\sigma}^2-\sigma^2)^2=$$
$$2E[(\hat{\sigma}^2-\hat{\sigma}^2_{EB})(\hat{\sigma}^2-\sigma^2)]-E(\hat{\sigma}^2-\hat{\sigma}^2_{EB})^2=$$
$$2J_2-J_1\tag{14}$$

由式(12)和式(8)可知：

$$\hat{\sigma}^2-\hat{\sigma}^2_{EB}=S^2-\frac{(m-1)S^2+m(\overline{X}-\hat{\theta})^2+\hat{\lambda}_n}{m+r+2}=$$
$$\frac{(r+3)S^2-m(\overline{X}-\hat{\theta})^2-\hat{\lambda}_n}{m+r+2}\tag{15}$$

从而

$$J_1=E(\hat{\sigma}^2-\hat{\sigma}^2_{EB})^2=$$
$$\frac{1}{(m+r+2)^2}E[(r+3)S^2-m(\overline{X}-\hat{\theta})^2-\hat{\lambda}_n]^2=$$
$$\frac{1}{(m+r+2)^2}[(r+3)^2ES^4+m^2E(\overline{X}-\hat{\theta})^4+E\hat{\lambda}_n^2-$$
$$2m(r+3)E[S^2(\overline{X}-\hat{\theta})^2]-2(r+3)E(S^2\hat{\lambda}_n)+2mE[(\overline{X}-\hat{\theta})^2\hat{\lambda}_n]]=$$
$$\frac{[(r+3)^2J_{11}+m^2J_{12}+J_{13}-2(r+3)mJ_{14}-2(r+3)J_{15}+2mJ_{16}]}{(m+r+2)^2}\tag{16}$$

其中由引理 2.2②得

$$J_{11}=ES^4=\frac{(m+1)\lambda^2}{(m-1)(r-2)(r-4)}\tag{17}$$

由于

$$(\overline{X}-\hat{\theta})\mid\tilde{\sigma}^2\sim N\left(0,\frac{1}{m}\left(\sigma^2+\frac{1}{n^2}\sum_{i=1}^n\sigma_i^2\right)\right),$$

记 $\sigma_*^2=\sigma^2+\frac{1}{n^2}\sum_{i=1}^n\sigma_i^2$，$Z=\frac{\sqrt{m}(\overline{X}-\hat{\theta})}{\sigma_*}$，则有 $Z\mid\tilde{\sigma}^2\sim N(0,1)$. 因此由引理 2.1③ 有

$$J_{12}=E(\overline{X}-\hat{\theta})^4=E[E[(\overline{X}-\hat{\theta})^4\mid\tilde{\sigma}^2]]=$$
$$E\left[\frac{\sigma_*^4}{m^2}E[Z^4\mid\tilde{\sigma}^2]\right]=$$
$$E\left[\frac{3\sigma_*^4}{m^2}\right]=\frac{3}{m^2}E\left(\sigma^2+\frac{1}{n^2}\sum_{i=1}^n\sigma_i^2\right)^2=$$
$$\frac{3\lambda^2[(r-2)n^3+2(r-4)n^2+(r-4)n+2]}{m^2n^3(r-2)^2(r-4)}\tag{18}$$

由引理 2.2③得

$$J_{13}=E(\hat{\lambda}_n^2)=\frac{(r-2)^2}{n^2}E\left(\sum_{i=1}^n S_i^2\right)^2=$$
$$\frac{(r-2)^2}{n^2}\cdot\frac{n\lambda^2[n(m-1)(r-4)+2r+2m-6]}{(m-1)(r-2)^2(r-4)}=$$
$$\frac{\lambda^2[n(m-1)(r-4)+2r+2m-6]}{n(m-1)(r-4)}\tag{19}$$

利用 $S^2\mid\tilde{\sigma}^2$ 和 $(\overline{X}-\hat{\theta})\mid\tilde{\sigma}^2$ 的条件独立性以及引理 2.1①，可得

$$J_{14}=E[E(S^2\mid\tilde{\sigma}^2)E((\overline{X}-\hat{\theta})^2\mid\tilde{\sigma}^2)]=$$
$$E\left[\sigma^2\cdot\frac{1}{m}\left(\sigma^2+\frac{1}{n^2}\sum_{i=1}^n\sigma_i^2\right)\right]=$$

$$\frac{\lambda^2[n(r-2)+r-4]}{mn(r-2)^2(r-4)} \tag{20}$$

由 S^2 和 $\hat{\lambda}_n$ 的独立性,可得

$$J_{15}=E(S^2)E(\hat{\lambda}_n)=\frac{\lambda}{r-2}\cdot\lambda=\frac{\lambda^2}{r-2} \tag{21}$$

由于 $(\overline{X}-\hat{\theta})\mid\tilde{\sigma}^2$ 和 $\hat{\lambda}_n\mid\tilde{\sigma}^2$ 条件独立以及引理2.1①,可得

$$\begin{aligned}J_{16}&=E[E[(\overline{X}-\hat{\theta})^2\mid\tilde{\sigma}^2]E[\hat{\lambda}_n\mid\tilde{\sigma}^2]]=\\&E\left[\frac{\sigma_*^2}{m}\cdot\frac{r-2}{n}\sum_{i=1}^{n}\sigma_i^2\right]=\\&\frac{r-2}{mn}E\left[\left(\sigma^2+\frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2\right)\sum_{i=1}^{n}\sigma_i^2\right]=\\&\frac{\lambda^2[n^2(r-4)+n(r-4)+2]}{mn^2(r-2)(r-4)}\end{aligned} \tag{22}$$

将式(17)~(22)代入式(16),并进行简单的化简:

$$\begin{aligned}J_1=&\frac{\lambda^2}{(m+r+2)^2(r-2)(r-4)}\cdot\\&(A\cdot r^2+B\cdot r+C+D)\end{aligned} \tag{23}$$

其中

$$A=\frac{m+1}{m-1}-1+\frac{2}{n(m-1)}=\frac{2(n+1)}{n(m-1)},$$

$$\begin{aligned}B=&\frac{6(m+1)}{m-1}-\frac{12}{n(m-1)}+\frac{2m+2}{n(m-1)}-\\&\frac{2}{n}+\frac{2}{n}-4=\\&-2+\frac{12}{m-1}-\frac{8}{n(m-1)}+\frac{2}{n},\end{aligned}$$

$$\begin{aligned}C=&\frac{9(m+1)}{m-1}+\frac{6}{n}-\frac{3}{n^2}+\frac{16}{n(m-1)}-\\&\frac{(4m+4)}{n(m-1)}-\frac{2}{n}-\frac{8}{n}+\frac{4}{n^2}+21=\\&30+\frac{18}{m-1}-\frac{8}{n}+\frac{7}{n^2}+\frac{8}{n(m-1)},\end{aligned}$$

$$\begin{aligned}D=&\frac{20}{n(r-2)}-\frac{12n^2+6n-6}{n^3(r-2)}=\\&\frac{2}{r-2}\left(\frac{4}{n}-\frac{3n-3}{n^3}\right).\end{aligned}$$

$$\begin{aligned}J_2=&E[(\hat{\sigma}^2-\hat{\sigma}_{EB})(\hat{\sigma}^2-\sigma^2)]=\\&E\left(\frac{(r+3)S^2-m(\overline{X}-\hat{\theta})^2-\hat{\lambda}_n}{m+r+2}\right)(S^2-\sigma^2)=\\&\frac{1}{m+r+2}[(r+3)E[S^2(S^2-\sigma^2)]-\\&mE[(\overline{X}-\hat{\theta})^2(S^2-\sigma^2)]-E[\hat{\lambda}_n(S^2-\sigma^2)]]=\\&\frac{1}{m+r+2}[(r+3)E(S^4-S^2\sigma^2)-0-0]=\\&\frac{2(r+3)\lambda^2}{(m+r+2)(m-1)(r-2)(r-4)}\end{aligned} \tag{24}$$

注 2.1 倒数第二个等号是因为:$(S^2-\sigma^2)\mid\tilde{\sigma}^2$ 与 $(\overline{X}-\hat{\theta})\mid\tilde{\sigma}^2$ 条件独立,$E[(S^2-\sigma^2)\mid\tilde{\sigma}^2]=0$,故第二项为0;$\hat{\lambda}_n$ 与 $S^2-\sigma^2$ 独立,$E(S^2-\sigma^2)=0$,故第三项也为0.

再将式(23)和式(24)代入式(14),我们可以得到

$$\begin{aligned}&MSE(\hat{\sigma}^2)-MSE(\hat{\sigma}^2_{EB})=2J_2-J_1=\\&\frac{\lambda^2}{(m+r+2)^2(r-2)(r-4)}\cdot\\&\left[\frac{4(r+3)(m+r+2)}{m-1}-(A\cdot r^2+B\cdot r+C+D)\right]=\\&\frac{\lambda^2}{(m+r+2)^2(r-2)(r-4)}\cdot\\&\left[\frac{2(n-1)}{n(m-1)}r^2+\left(2+\frac{12}{m-1}+\frac{8}{n(m-1)}-\frac{2}{n}\right)\cdot r+\right.\\&\left(-18+\frac{18}{m-1}+\frac{8}{n}-\frac{7}{n^2}-\frac{8}{n(m-1)}\right)+\\&\left.\frac{2}{r-2}\left(\frac{-4}{n}+\frac{3n-3}{n^3}\right)\right]=\\&\frac{\lambda^2}{(m+r+2)^2(r-2)(r-4)}(A_1+B_1\cdot r+C_1)\end{aligned} \tag{25}$$

其中

$$A_1=2r-18,$$

$$B_1=\frac{2(n-1)}{n(m-1)}r+\frac{12}{m-1}-\frac{2}{n},$$

$$\begin{aligned}C_1=&\frac{8}{n(m-1)}(r-1)+\frac{18}{m-1}+\frac{8}{n}-\\&\frac{7}{n^2}-\frac{8}{n(r-2)}+\frac{6(n-1)}{n^3(r-2)}.\end{aligned}$$

当 $r\geqslant 9$ 时,$A_1\geqslant 0$;

$r\geqslant 9,n\geqslant 2\Rightarrow\frac{n-1}{n}\geqslant\frac{1}{2}\Rightarrow\frac{2(n-1)}{n(m-1)}r+\frac{12}{m-1}-\frac{2}{n}>\frac{21}{m}-\frac{2}{n}$,故当 $n>\frac{m}{10.5}$ 时,$B_1>0$;

$r\geqslant 9,n\geqslant 2,m>1\Rightarrow\frac{8}{n}-\frac{7}{n^2}-\frac{8}{n(r-2)}\geqslant\frac{8}{n}-\frac{7}{2n}-\frac{8}{7n}>0$,$\frac{8}{n(m-1)}(r-1)$,$\frac{18}{m-1}$ 与 $\frac{6(n-1)}{n^3(r-2)}$ 俱大于0,$C_1>0$.

因此当 $r\geqslant 9,n>\max\left(1,\frac{m}{10.5}\right),m>1$ 时,有式(25)是恒大于零的.

到此我们证明了定理2.1中的式(13):

$\mathrm{MSE}(\hat{\sigma}^2)-\mathrm{MSE}(\hat{\sigma}^2_{EB})>0$.

3　超参数 r 和 λ 皆未知时,PEB 估计及其优良性

本节将在 r 和 λ 皆未知的情形下,构造 σ^2 的 PEB 估计. 由于此时 σ^2 的 PEB 估计比前一节的复杂,r 的估计出现在分母上,在 MSE 准则下的小样本性质的理论结果很难获得,故本节我们通过数值模拟比较 σ^2 的 PEB 估计和 UMVUE 的均方误差,获得其优良性.

3.1　PEB 估计的构造

当超参数 r 和 λ 皆未知时,需要利用历史样本导出 r 和 λ 的估计量,用它们代替式(7)中的 r 和 λ,从而获得 σ^2 的 PEB 估计. 我们采用矩估计方法:

$$ES^2=\frac{\lambda}{r-2} \tag{26}$$

$$ES^4=\frac{m+1}{m-1}\cdot\frac{\lambda^2}{(r-2)(r-4)} \tag{27}$$

解方程组得:

$$r=\frac{4ES^4(m-1)-2(ES^2)^2(m+1)}{ES^4(m-1)-(ES^2)^2(m+1)} \tag{28}$$

$$\lambda=ES^2\cdot\frac{2ES^4(m-1)}{ES^4(m-1)-(ES^2)^2(m+1)} \tag{29}$$

用 $\overline{S}^2=\frac{1}{n}\sum_{i=1}^{n}S_i^2$ 和 $\overline{S}^4=\frac{1}{n}\sum_{i=1}^{n}S_i^4$ 分别代替 ES^2 和 ES^4,得 r 和 λ 的矩估计量:

$$\left.\begin{aligned}\tilde{r}_n&=\frac{4\overline{S}^4(m-1)-2(\overline{S}^2)^2(m+1)}{\overline{S}^4(m-1)-(\overline{S}^2)^2(m+1)},\\ \tilde{\lambda}_n&=\overline{S}^2\cdot\frac{2\overline{S}^4(m-1)}{\overline{S}^4(m-1)-(\overline{S}^2)^2(m+1)}\end{aligned}\right\} \tag{30}$$

$\hat{\theta}$ 仍由式(9)给出,将式(7)中 θ, r 和 λ 用 $\hat{\theta}$, $\tilde{r}_n$ 和 $\tilde{\lambda}_n$ 代入,得 σ^2 的 PEB 估计为

$$\hat{\sigma}^2_{EB}=\frac{1}{m+\tilde{r}_n+2}[(m-1)S^2+m(\overline{X}-\hat{\theta})^2+\tilde{\lambda}_n] \tag{31}$$

3.2　PEB 估计在 MSE 准则下优良性的模拟结果

我们将采用模拟的方法,比较此种情形下 σ^2 的 PEB 估计 $\hat{\sigma}^2_{EB}$ 与 UMVUE $\hat{\sigma}^2$ 的优良性. 我们的方法是对于每次模拟,让历史样本个数,θ,λ,r 其中的一个量发生变化,其他量固定不变. 我们从误差方差的先验分布 $\Gamma^{-1}(r/2,\lambda/2)$ 中抽取出 $\sigma_1^2,\cdots,\sigma_n^2,\sigma^2$. 然后从 $N(\theta,\sigma_i^2)$ 中产生样本 $\mathbf{X}^{(i)}=(X_{i1},\cdots,X_{im})$, $i=1,2\cdots,n$ 作为历史样本;从 $N(\theta,\sigma^2)$ 中产生样本 $\mathbf{X}^{(n+1)}=\mathbf{X}=(X_1,\cdots,X_m)$ 作为当前样本. 当前样本用于计算 UMVUE,从而得出 $(\hat{\sigma}^2-\sigma^2)^2$;历史样本用于估计未知参数,再联合当前样本计算 PEB 估计,从而得出 $(\hat{\sigma}^2_{EB}-\sigma^2)^2$,这样不断重复 1 000 次,由蒙特卡罗法得到数值计算的 MSE. 模拟结果如图 1 所示.

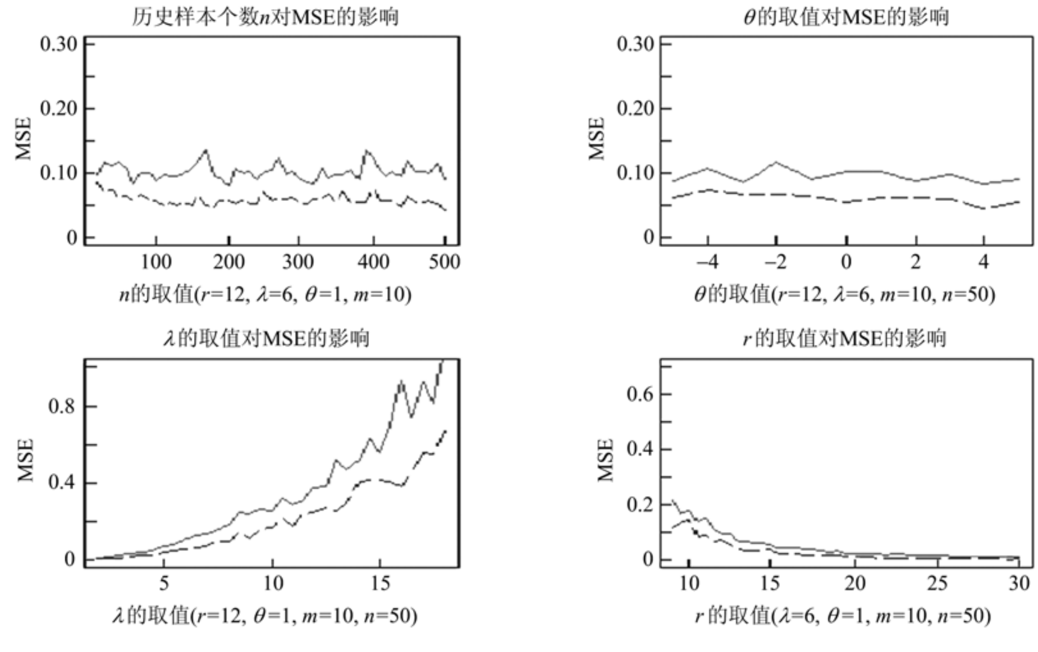

图 1　σ^2 的 UMVUE(实线)和 PEB(虚线)的 MSE 比较

Fig. 1　The MSE of UMVUE (solid) of σ^2 and the MSE of PEB (dotted) of σ^2

由图 1 可得如下结论：

（Ⅰ）历史样本个数对于两者的影响都不大，PEB 估计一直优于 UMVUE，做估计时，拥有一定数量的历史样本就够了，控制了数据成本；

（Ⅱ）θ 的变化对两者的影响都很小，PEB 估计一直优于 UMVUE；

（Ⅲ）对所有的 λ，PEB 估计一直优于 UMVUE，随着 λ 的增加，两者的 MSE 都发散变大；

（Ⅳ）r（$r\geqslant 9$）较小时，PEB 估计优于 UMVUE，但是随着 r 的增加，两者的 MSE 都迅速地降低并趋于一样.

致谢 感谢韦来生教授对本文的指导.

参考文献(References)

[1] Box G E P, Tiao G C. Bayesian Inference in Statistical Analysis[M]. MA: Reading, 1973.

[2] Mao Shisong. Bayes Statistics[M]. Beijing: Statistics Press of China, 1999.
茆诗松. 贝叶斯统计[M]. 北京：中国统计出版社，1999.

[3] Morris C. Parametric empirical Bayes inference: Theory and applications[J]. Journal of the American Statistical Association, 1983, 78: 47-65.

[4] Singh R S. Empirical Bayes estimation in lebesgue-exponential families with rates near the best possible rate[J]. The Annals of Statistics, 1979, 7(4): 890-902.

[5] Wei Laisheng. Convergence rates of empirical estimation for the parameter in continuous multi-parameter exponential family[J]. Acta Mathematica Sinica, 1987, 30(2): 272-279.
韦来生. 连续多参数指数族参数的经验 Bayes 估计的收敛速度[J]. 数学学报，1987，30(2)：272-279.

[6] Wang Lichun, Wei Laisheng. Convergence rates of empirical estimation for the parameter in scale-exponential family [J]. Chinese Annals of Mathematics, 2002, 23A(5): 555-564.
王立春，韦来生. 刻度指数族参数的经验 Bayes 估计的收敛速度[J]. 数学年刊，2002，23A(5)：555-564.

[7] Berger J O. Statistical Decision Theory and Bayesian Analysis[M]. Berlin: Springer, 1985.

[8] Zhang Weiping, Wei Laisheng, Yang Yaning. The superiorities of empirical Bayes estimator of parameters in linear model[J]. Statistics and Probability Letters, 2005, 72: 43-50.

[9] Zhang Weiping, Wei Laisheng. The superiority of empirical Bayes estimation of parameters in partitioned normal linear model[J]. Acta Mathematica Scientia, 2008, 28B(4): 955-962.

[10] Wei Laisheng. Mathematical Statistics[M]. Beijing: Science Press, 2008: 312-313.
韦来生. 数理统计[M]. 北京：科学出版社，2008：312-313.