DOI:10.11721/cqnuj20170202

NA 样本下艾拉姆咖分布参数的经验 Bayes 检验*

张　月，周菊玲
(新疆师范大学 数学科学学院，乌鲁木齐 830017)

摘要：【目的】研究 NA 样本下艾拉姆咖分布参数的经验 Bayes 检验问题。【方法】在同分布负相协(NA)随机列$\{X_1,X_2,\cdots,X_n\}$下，利用概率密度函数的变窗核估计方法，讨论了艾拉姆咖分布参数 θ 的经验 Bayes 检验问题。【结果】首先得到了经验 Bayes 检验函数 $\delta_n(x)$，然后证明了 $\delta_n(x)$的渐近最优性。【结论】在适当的条件下，利用相关引理和不等式，可获得参数 θ 的经验 Bayes 检验函数 $\delta_n(x)$的收敛速度为 $O(n^{-\frac{1}{2}})$。

关键词：NA 样本；艾拉姆咖分布；经验 Bayes 检验；渐近最优性
中图分类号：O212.1　**文献标志码：**A　**文章编号：**1672-6693(2017)02-0049-04

在研究武器装备的维修时间时，俄罗斯首先引入了艾拉姆咖分布，该分布对装备维修理论的研究起着重要作用。对此分布统计性质的研究，国内较多集中在样本相互独立情形下，比如：讨论了艾拉姆咖分布的极大似然估计[1]；研究了艾拉姆咖分布在小样本下的区间估计和假设检验问题[2]；在全样本情形下讨论了艾拉姆咖分布的 Bayes 估计[3]。还有一些文献，讨论了艾拉姆咖分布参数在不同先验分布下，截尾场合下及其次序统计量的估计[4-5]，但是也都在简单随机样本下，对相依样本讨论的极少。本文则是在负相协(NA)样本情形下就该分布进行了讨论，得到了其中参数 θ 的经验 Bayes 检验函数，并获得 Bayes 检验函数的渐近性质和收敛速度。

下面首先引入 NA 随机变量序列的定义：

定义 1[6]　称 r.v. $X_1,X_2,\cdots,X_n$ 为负相协(NA)的，如果对于$\{1,2,\cdots,n\}$的任何两个不交的非空子集 A_1 和 A_2 都有 $Cov(f_1(X_i,i\in A_1),f_2(X_j,j\in A_2))\leqslant 0$，其中 f_1 及 f_2 是任何两个使得协方差存在且对每个变元均非降(或非增)的函数。称 r.v.序列$\{X_j,j\in\mathbf{N}\}$是 NA 的，如果对任何自然数 $n\geqslant 2$，$X_1,X_2,\cdots,X_n$ 都是 NA 的。

定义 2　称 r.v. X 服从艾拉姆咖分布，如果其概率函数为 $p(x;\theta)=\frac{4x}{\theta^2}\mathrm{e}^{-\frac{2x}{\theta}},x>0$。其中 $\theta>0$ 为装备维修效率。记样本空间 $\aleph=(0,\infty)$，参数空间 $\Theta=(0,\infty)$。

1 经验 Bayes 检验函数的构造

考虑艾拉姆咖分布中参数 θ 的如下经验 Bayes 假设检验问题：

$$H_0:\theta\leqslant\theta_0 \text{ vs } H_1:\theta>\theta_0$$

设损失函数为 $L_i(\theta,d_i)=(1-i)\left(1-\frac{\theta_0}{\theta}\right)I_{[\theta>\theta_0]}+i\left(\frac{\theta_0}{\theta}-1\right)I_{[\theta\leqslant\theta_0]},i=0,1$。其中 θ_0 为已知正常数，d_i 表示接受 H_i，$D=\{d_0,d_1\}$ 为行动空间，$I_{[A]}$为示性函数。

设参数 θ 的先验分布为$\pi(\theta)$。在 Bayes 统计中，用 x 在 θ 下的条件概率函数 $p(x|\theta)$表示总体概率函数[7]。

设随机化判决函数为 $\delta(x)=P(d_0|X=x)$，则在 $\pi(\theta)$下，X 的边缘分布为

$$m(x)=\int_\Theta p(x|\theta)\mathrm{d}\pi(\theta)=\int_\Theta\frac{4x}{\theta^2}e^{-\frac{2x}{\theta}}\mathrm{d}\pi(\theta)。$$

$\delta(x)$ 的风险函数为

$$R(\delta(x),\pi(\theta))=\int_\Theta\int_\aleph[L_0(\theta,d_0)\delta(x)+L_1(\theta,d_1)(1-\delta(x))]p(x|\theta)\mathrm{d}x\mathrm{d}\pi(\theta)=\int_\aleph\alpha(x)\delta(x)\mathrm{d}x+C_\pi,$$

*　**收稿日期：**2016-04-01　**修回日期：**2016-04-27　**网络出版时间：**2017-03-13 11:08
第一作者简介：张月，女，研究方向为概率论与数理统计，E-mail：1589663192@qq.com；**通信作者：**周菊玲，副教授，E-mail：326815649@qq.com
网络出版地址：http://kns.cnki.net/kcms/detail/50.1165.N.20170313.1108.048.html

其中 $\alpha(x)=\int_{\Theta}\left(1-\frac{\theta_0}{\theta}\right)p(x\mid\theta)\mathrm{d}\pi(\theta)=\left(1-\frac{\theta_0}{2x}\right)m(x)+\frac{\theta_0}{2}m'(x)$，$C_\pi=\int_{\Theta}L_1(\theta,d_1)\mathrm{d}\pi(\theta)$。

令 $u(x)=1-\frac{\theta_0}{2x}$，$v(x)=\frac{\theta_0}{2}$，则 $\alpha(x)=u(x)m(x)+v(x)m'(x)$，故 Bayes 检验函数为 $\delta_\pi(x)=\begin{cases}1,\alpha(x)\leqslant 0\\0,\alpha(x)>0\end{cases}$。

Bayes 风险函数为：

$$R(\pi)=\inf_{\delta}R(\delta(x),\pi(\theta))=R(\delta_\pi,\theta)=\int_{\aleph}\alpha(x)\delta_\pi(x)\mathrm{d}x+C_\pi。\tag{1}$$

因为此处先验分布 $\pi(\theta)$ 是未知的，所以 $\delta_\pi(x)$ 无使用价值，故而引入经验 Bayes 方法。

设 r.v. 序列 $\{X_1,X_2,\cdots,X_n,X\}$ 是同分布弱平稳 NA 的，其中 $X_1,X_2,\cdots,X_n$ 是历史样本，X 是当前样本。参数 θ 的先验分布为 $\pi(x)$。本文假定：

a) 对于 X 的边缘密度函数为 $m(x)$，有 $m(x)\in C_{s,a}$，$x\in\mathbf{R}^1$，其中 $C_{s,a}$ 表示 s 阶连续可导且绝对值不超过正常数 a 的概率密度族，$s\geqslant 3$ 为正整数；

b) NA 随机样本序列的协方差满足 $\sum_{j=1}^{n}|Cov(X_j,X)|\leqslant c<\infty$；

c) 设 $k_i(\cdot)$ 是 $(0,1)$ 外为零，$(0,1)$ 内有界的 Borel 可测函数，$i=0,1$。满足条件：

c$_1$) $\frac{1}{t!}\int_0^1 x^t k_i(x)\mathrm{d}x=\begin{cases}1,\ t=i\\0,\ t\neq i,\ t=1,2,\cdots,s-1;\end{cases}$

c$_2$) $k_i(x)$ 在 $\mathbf{R}^1$ 上几乎处处可微，且 $\sup_{x\in\mathbf{R}^1}|k_i'(x)|\leqslant c<\infty$。

定义 3[8] $m^{(i)}(x)$ 的变窗核估计为 $m_n^{(i)}(x)=\frac{1}{n}\sum_{j=1}^{n}\frac{1}{b_j^{1+i}}k_i\left(\frac{X_j-x}{b_j}\right)$，$i=0,1$；其中 b_j 为窗宽且 $b_j\to 0$。则 $\alpha_n(x)=u(x)m_n(x)+v(x)m_n^{(1)}(x)$。

那么可定义 θ 的经验 Bayes 检验函数为 $\delta_n(x)=\begin{cases}1,\alpha_n(x)\leqslant 0\\0,\alpha_n(x)>0\end{cases}$。

设 E_n 表示对 $X_1,X_2,\cdots,X_n$ 联合分布求期望。所以 $\delta_n(x)$ 的全面 Bayes 风险函数为

$$R(\delta_n,\theta)=E_n\left[\int_{\aleph}\alpha(x)\delta_n(x)\mathrm{d}x+C_\pi\right]=\int_{\aleph}\alpha(x)E_n\delta_n(x)\mathrm{d}x+C_\pi。\tag{2}$$

定义 4 设 F 表示参数 θ 的先验分布族。若 $\forall\pi(\theta)\in F$，有 $\lim_{n\to\infty}R(\delta_n,\theta)=R(\pi)$，称 $\{\delta_n(x)\}$ 为渐近最优的经验 Bayes 检验函数。若 $R(\delta_n,\theta)-R(\pi)=O(n^{-q})$，$q>0$，则称经验 Bayes 检验函数具有收敛速度的阶为 $O(n^{-q})$。

为导出 $\{\delta_n(x)\}$ 的渐近最优性和收敛速度的阶，需引理 1～3。下文中 M,M_1,M_2 均表示正常数；h,H 分别表示最小窗宽和最大窗宽。

引理 1[9] 设 $R(\pi)$、$R(\delta_n,\theta)$ 为(1),(2)式定义，则

$$0\leqslant R(\delta_n,\theta)-R(\pi)\leqslant\int_{\aleph}|\alpha(x)|\cdot P(|\alpha_n(x)-\alpha(x)|\geqslant|\alpha(x)|)\mathrm{d}x。$$

引理 2[10] 设 X 和 Y 是 NA 随机变量且方差有限，则对于任何可微函数 $g_1(x)$，$g_2(y)$ 有

$$|\mathrm{Cov}(g_1(x),g_2(y))|\leqslant\sup_x|g_1'(x)|\cdot\sup_y|g_2'(y)|\cdot|\mathrm{Cov}(X,Y)|。$$

引理 3 设 r.v. 序列 $\{X_1,X_2,\cdots,X_n\}$ 是同分布弱平稳 NA 的，$m_n^{(i)}(x)$，$i=0,1$ 由定义 3 给出，a)、b)、c) 成立，则当 $H\to 0$，$(nh^6)^{-1}\to 0(n\to\infty)$ 时，有 $E_n|m_n^{(i)}(x)-m^{(i)}(x)|\to 0(n\to\infty)$。

证明 由 C_r 不等式知，$E_n|m_n(x)-m(x)|\leqslant M_1|E_nm_n(x)-m(x)|+M_2|\mathrm{Var}\ m_n(x)|^{\frac{1}{2}}$，

因为 $E_nm_n(x)=\frac{1}{n}\sum_{j=1}^{n}\int_0^1 k_0(u)m(x+ub_j)\mathrm{d}u$，将 $m(x+ub_j)$ 在 x 处泰勒展开有

$$m(x+ub_j)=m(x)+\frac{m^{(1)}(x)}{1!}ub_j+\cdots+\frac{m^{(s)}(x+\beta_jub_j)}{s!}\cdot(ub_j)^s,0<\beta_j<1。$$

再由假设条件 a) 和 c$_1$) 可得

$$E_nm_n(x)=\frac{1}{n}\sum_{j=1}^{n}\int_0^1 k_0(u)\left[m(x)+m^{(1)}(x)(ub_j)+\cdots+\frac{m^{(s)}(x+\beta_jub_j)}{s!}(ub_j)^s\right]\mathrm{d}u=$$

$$m(x)+\frac{1}{n}\sum_{j=1}^{n}\int_{0}^{1}k_0(u)\cdot\frac{m^{(s)}(x+\beta_j ub_j)}{s!}\cdot(ub_j)^s\mathrm{d}u,$$

从而有 $|E_n m_n(x)-m(x)|\leqslant\frac{1}{n}\sum_{j=1}^{n}\int_{0}^{1}|k_0(u)|\cdot\left|\frac{m^{(s)}(x+\beta_j ub_j)}{s!}\right|\cdot u^s b_j^s\mathrm{d}u\leqslant MH^s$，

另一方面　$\mathrm{Var}[m_n(x)]=\mathrm{Var}\left[\frac{1}{n}\sum_{j=1}^{n}\frac{1}{b_j}k_0\left(\frac{X_j-x}{b_j}\right)\right]=$

$$\frac{1}{n^2}\sum_{j=1}^{n}\mathrm{Var}\left[\frac{1}{b_j}k_0\left(\frac{X_j-x}{b_j}\right)\right]+\frac{2}{n^2}\sum_{1\leqslant j<k\leqslant n}\mathrm{Cov}\left(\frac{1}{b_j}k_0\left(\frac{X_j-x}{b_j}\right),\frac{1}{b_k}k_0\left(\frac{X_k-x}{b_k}\right)\right)\triangleq D_1+D_2,$$

则 $D_1\leqslant\frac{1}{n^2}\sum_{j=1}^{n}E_n\left[\frac{1}{b_j}k_0\left(\frac{X_j-x}{b_j}\right)\right]^2=\frac{1}{n^2}\sum_{j=1}^{n}\frac{1}{b_j}\int_{0}^{1}k_0^2(u)\cdot m(x+ub_j)\mathrm{d}u\leqslant\frac{M}{nh}$。

令 $g_j(X_j,x)=k_0\left(\frac{X_j-x}{b_j}\right)$，由引理 2 及假设条件 c_2)，有

$$D_2=\frac{2}{n^2}\sum_{1\leqslant j<k\leqslant n}\frac{1}{b_jb_k}\mathrm{Cov}\left(k_0\left(\frac{X_j-x}{b_j}\right),k_0\left(\frac{X_k-x}{b_k}\right)\right)\leqslant$$

$$\frac{2}{n^2}\sum_{1\leqslant j<k\leqslant n}\frac{1}{b_jb_k}|\sup g_j'(X_j-x)|\cdot|\sup g_k'(X_k-x)|\cdot|\mathrm{Cov}(X_j,X_k)|\leqslant$$

$$\frac{2}{n^2h^2}\sum_{1\leqslant j<k\leqslant n}\frac{1}{b_jb_k}|\sup k_0'(x)|\cdot|\sup k_0'(x)|\cdot|\mathrm{Cov}(X_j,X_k)|\leqslant\frac{2M_1}{nh^4}\sum_{1\leqslant j<k\leqslant n}|\mathrm{Cov}(X_j,X_k)|\leqslant\frac{M}{nh^4},$$

综上，

$$|E_n m_n(x)-m(x)|\leqslant MH^s,\ \mathrm{Var}[m_n(x)]\leqslant\frac{M}{nh^4},\tag{3}$$

当 $h<H\to0,(nh^4)^{-1}\to0$ 时，$E_n|m_n(x)-m(x)|\to0,(n\to\infty)$。类似的有结果：

$$|E_n m_n^{(1)}(x)-m^{(1)}(x)|\leqslant MH^{s-1},\ \mathrm{Var}[m_n(x)]\leqslant\frac{M}{nh^6},\tag{4}$$

当 $H\to0,(nh^6)^{-1}\to0$ 时，$E_n|m_n^{(1)}(x)-m^{(1)}(x)|\to0,(n\to\infty)$。　证毕

2 主要结果及证明

定理 1　设 r. v. 序列 $\{X_1,X_2,\cdots,X_n\}$ 是同分布弱平稳 NA 的。$R(\pi)$，$R(\delta_n,\theta)$ 如(1)，(2)式定义，假定 a)，b)，c)成立，则当 $H\to0,(nh^6)^{-1}\to0(n\to\infty)$ 时，$\lim\limits_{n\to\infty}R(\delta_n,\theta)=R(\pi)$。

证明　$|\alpha(x)|=\int_{\Theta}\left|\left(1-\frac{\theta_0}{\theta}\right)p(x|m)\right|\mathrm{d}\pi(\theta)<\infty$，由引理 1 和控制收敛定理有

$$0\leqslant\lim_{n\to\infty}(R(\delta_n,\theta)-R(\pi))\leqslant\int_{\mathfrak{X}}\lim_{n\to\infty}|\alpha(x)|\cdot P(|\alpha_n(x)-\alpha(x)|\geqslant|\alpha(x)|),$$

由 Markov 不等式

$$|\alpha(x)|\cdot P(|\alpha_n(x)-\alpha(x)|\geqslant|\alpha(x)|)\leqslant E_n|\alpha_n(x)-\alpha(x)|\leqslant$$

$$|u(x)|\cdot E_n|m_n(x)-m(x)|+|v(x)|\cdot E_n(x)|m_n^{(1)}(x)-m^{(1)}(x)|。$$

由引理 3，当 $n\to\infty$ 时，$E_n|m_n^{(i)}(x)-m^{(i)}(x)|\to0,i=0,1$。所以 $\lim\limits_{n\to\infty}R(\delta_n,\theta)=R(\pi)$。也就是说经验 Bayes 检验函数 $\{\delta_n(x)\}$ 是渐近最优的。　证毕

定理 2　设 r. v. 序列 $\{X_1,X_2,\cdots,X_n\}$ 是同分布弱平稳 NA 的。$R(\pi)$，$R(\delta_n,\theta)$ 如(1)，(2)式定义，$m_n^{(i)}(x)$，$i=0,1$ 由定义 3 给出，假定 a)，b)，c) 成立，对于 $\forall\beta\in(0,1)$，满足 $\int_{\mathfrak{X}}|\alpha(x)|^{1-\beta}|u(x)|^\beta\mathrm{d}x<\infty$，$\int_{\mathfrak{X}}|\alpha(x)|^{1-\beta}|v(x)|^\beta\mathrm{d}x<\infty$ 成立，则当 $n^{\frac{1}{2s+1}}<h<H<n^{\frac{2}{2s+3}}$ 时，$R(\delta_n,\pi)-R(\pi)=O(n^{\frac{\beta(2s-5)}{4s+2}})$。

证明　由引理 1 和 Markov 不等式得

$$0\leqslant R(\delta_n,\pi)-R(\pi)\leqslant\int_{\mathfrak{X}}|\alpha(x)|\cdot P(|\alpha_n(x)-\alpha_\pi(x)|\geqslant|\alpha(x)|)\mathrm{d}x\leqslant\int_{\mathfrak{X}}|\alpha(x)|^{1-\beta}E_n|\alpha_n(x)-\alpha_\pi(x)|^\beta\mathrm{d}x\leqslant$$

$$\int_{\mathfrak{X}}|\alpha(x)|^{1-\beta}|u(x)|^\beta E_n|m_n(x)-m(x)|^\beta dx+\int_{\mathfrak{X}}|\alpha(x)|^{1-\beta}|v(x)|^\beta E_n|m_n^{(1)}(x)-m^{(1)}(x)|^\beta\mathrm{d}x。$$

当 $n^{\frac{1}{2s+1}}<h<H<n^{\frac{2}{2s+3}}$，$s\geqslant 3$ 时，由(3)，(4)式得：

$$E_n\left|m_n(x)-m(x)\right|^{\beta}\leqslant Mn^{\frac{\beta(2s-3)}{4s+2}},E_n\left|m_n^{(1)}(x)-m^{(1)}(x)\right|^{\beta}\leqslant Mn^{\frac{\beta(2s-5)}{4s+2}}。$$

所以 $0\leqslant R(\delta_n,\pi)-R(\pi)\leqslant Mn^{-\frac{\beta(2s-5)}{4s+2}}$。 证毕

参考文献：

[1] 吕会强，高连华，陈春良. 艾拉姆咖分布及其在保障性数据分析中的应用[J]. 装甲兵工程学院学报，2002，16(3)：48-52.
LYU H Q，GAO L H，CHEN C L. ЭРланга distribution and its application in supportability data analysis[J]. Journal of Academy of Armored Forces Engineering，2002，16(3)：48-52.

[2] 潘高田，王保恒，陈春良，等. 艾拉姆咖分布小样本区间估计和假设检验问题研究[J]. 数理统计与管理，2009，28(3)：468-472.
PAN G T，WANG B H，CHEN C L，et al. The research of interval estimation and hypothetical test of small sample of ЭРланга distribution[J]. Application of Statistics and Management，2009，28(3)：468-472.

[3] 龙兵. 不同损失函数下艾拉姆咖分布参数的 Bayes 估计—全样本情形[J]. 重庆师范大学学报(自然科学版)，2013，30(5)：96-99.
LONG B. Bayesian estimation of ЭРланга distribution parameter under different loss function-all samples[J]. Journal of Chongqing Normal University (Natural Science)，2013，30(5)：96-99.

[4] 龙兵. 不同先验分布下艾拉姆咖分布参数的 Bayes 估计[J]. 数学的实践与认识，2015，45(4)：186-191.
LONG B. Bayesian estimation of ЭРланга distribution parameter under different loss function[J]. Mathematics in Practice and Theory，2015，45(4)：186-191.

[5] 姜培华. 艾拉姆咖分布顺序统计量的概率分布及渐近性质[J]. 南通大学学报(自然科学版)，2015，14(2)：64-68.
JIANG P H. Probability distribution and asymptotic properties of order statistics from ЭРланга distribution[J]. Journal of Nantong Normal University (Natural Science)，2015，14(2)：64-68.

[6] JOAGDEV K，PROSCHAN F. Negative association of random variables with application[J]. Ann Statist，1983，11：286-195.

[7] 茆诗松，程依明，濮晓龙. 概率论与数理统计教程[M]. 北京：高等教育出版社，2011：334-335.
MAO S S，CHENG Y M，PU X L. Probability theory and mathematical statistics tutorial[M]. Beijing：Higher Education Press，2011：334-335.

[8] 王亮，师义民. NA 样本下一类指数分布族的经验贝叶斯检验[J]. 西北大学学报(自然科学版)，2008，38(4)：523-526.
WANG L，SHI Y M. The empirical Bayes test for one exponential family using NA samples[J]. Journal of Northwest University (Natural Science edition)，2008，38(4)：523-526.

[9] JOHNS M V J，VAN R J. Convergence rates in empirical Bayes two-action problems Ⅱ discrete case[J]. Ann Math Statist，1971，42：1521-1539.

[10] PAN J M. On the convergence rates in the central limit theorem for negatively associated sequences[J]. Chinese Journal of Applied Probability and Statistics，1997，13(2)：183-192.

[11] 周慧，阳连武. NA 样本下 Burr Ⅱ型分布形状参数的经验 Bayes 检验[J]. 统计与决策，2012(15)：7-9.
ZHOU H，YANG L W. The empirical Bayes test for Burr Ⅱ distribution using NA samples[J]. Statistics and Decision Making，2012(15)：7-9.

The Empirical Bayes Test Problem for ЭРланга Using NA Sample

ZHANG Yue，ZHOU Juling

(Department of Mathematical Sciences，Xinjiang Normal University，Urumqi 830017，China)

Abstract：[**Purposes**]To study the empirical Bayes test problem for ЭРланга under negatively associated using NA sample. [**Methods**] The empirical Bayes test for the parameter θ of ЭРланга distribution is discussed by using the kernel estimation method with variable window width of probability density function in the same distribution with NA. [**Findings**]The empirical Bayes test function $\delta_n(x)$ is obtained first，and then the asymptotic optimality of $\delta_n(x)$ is proved. [**Conclusions**]Using related lemmas and inequalities，convergence rate of the empirical Bayes test function of parameter θ is obtained under certain conditions，which can close to $O(n^{-\frac{1}{2}})$ arbitrarily.

Keywords：NA sample；ЭРланга distribution；empirical Bayes test；asymptotic optimality

（责任编辑 许 甲）