

不同损失函数下艾拉姆咖分布参数的 Bayes 估计* ——全样本情形

龙 兵

(荆楚理工学院 数理学院, 湖北 荆门 448000)

摘要:艾拉姆咖分布在装备的维修理论中具有重要的作用。文中首先给出了这类分布在全样本场合下参数的极大似然估计;由于参数的 Bayes 估计的优良性与所选取的损失函数有关,因此分别在熵损失、Linux 损失、二次损失、平方损失和平衡损失函数下,给出了参数的 Bayes 估计如 $2n / (2 \sum_{i=1}^n t_i + \lambda)$ 等,并证明了所给估计都是容许的;通过随机模拟比较了几类估计的均方误差以及与真值的偏差,从这两方面来看,在熵损失和平衡损失函数下 θ 的 Bayes 估计是较优的;最后通过一个实例,计算了几类估计的值并进行了分析,显示几类估计值相差不大。

关键词:艾拉姆咖分布;先验分布;损失函数;Bayes 估计

中图分类号:O212.8

文献标志码:A

文章编号:1672-6693(2013)05-0096-05

俄罗斯在研究武器装备的维修时间时引入了艾拉姆咖分布,此分布在装备的维修理论中具有重要的作用。国内对这类分布统计性质进行研究的文献很少。文献[1]对艾拉姆咖分布的特点进行了分析,其次在全样本场合下运用极大似然法对分布的参数进行了估计,并通过实例验证了这种分布的可行性和实用性。文献[2]研究了艾拉姆咖分布的小样本区间估计和检验问题,并运用实例指出在对装备维修工时的估计时,用艾拉姆咖分布进行估计的精度比用指数分布高。文献[3]在定数截尾样本下研究了参数的极大似然估计,并在全样本场合下给出了参数的精确区间估计和近似区间估计,最后用实例说明精确区间估计优于近似区间估计。然而,关于艾拉姆咖分布参数的 Bayes 估计的讨论尚未见到,本文将在不同损失函数下,给出了参数的 Bayes 估计,最后通过实例进行了分析。

艾拉姆咖分布的分布函数和密度函数分别为

$$F(t, \theta) = 1 - (1 + 2\theta t)e^{-2\theta t} \quad (1) \quad f(t, \theta) = 4\theta^2 te^{-2\theta t} \quad (2)$$

其中 $t > 0$, 参数 $\theta > 0$, 由于 $\theta = \frac{1}{E(T)}$, 则 θ 为装备的维修效率, 因此对参数的估计是很有意义的。

1 参数 θ 的极大似然估计

设 t_1, t_2, \dots, t_n 为来自艾拉姆咖分布容量为 n 的随机样本, 令 $t = (t_1, t_2, \dots, t_n)$, 则样本 t 的似然函数为

$$L(t | \theta) = \prod_{i=1}^n f(t_i, \theta) = 4^n \theta^{2n} \prod_{i=1}^n t_i e^{-2\theta \sum_{i=1}^n t_i} \quad (3)$$

$$\text{令 } \frac{\partial \ln L(t | \theta)}{\partial \theta} = \frac{2n}{\theta} - 2 \sum_{i=1}^n t_i = 0, \text{ 得到参数 } \theta \text{ 的极大似然估计为} \quad \hat{\theta}_{MLE} = \frac{n}{\sum_{i=1}^n t_i} \quad (4)$$

2 参数 θ 的 Bayes 估计

2.1 熵损失函数下 θ 的 Bayes 估计

定义 1^[4] 随机变量 X 服从密度函数为 $f(x, \theta)$ 的分布, 其中 θ 为参数, 如果 δ 是 θ 的判别空间的一个估计,

* 收稿日期:2012-11-01 修回日期:2013-02-27 网络出版时间:2013-09-17 17:38

资助项目:湖北省教育厅重点科研项目(No. D20134301)

作者简介:龙兵,男,副教授,硕士,研究方向为概率统计,E-mail:qh-longbing@163.com

网络出版地址:http://www.cnki.net/kcms/detail/50.1165.N.20130917.1738.201305.96_016.html

则熵损失函数定义为似然比对数的数学期望,即 $L(\theta, \delta) = E_{\theta} \left\{ \ln \frac{f(\theta, X)}{f(\delta, X)} \right\}$ 。根据上述定义,艾拉姆咖分布的熵损

失函数为 $L(\theta, \delta) = E_{\theta} \left\{ \ln \frac{4^n \theta^{2n} \prod_{i=1}^n t_i e^{-2\theta \sum_{i=1}^n t_i}}{4^n \delta^{2n} \prod_{i=1}^n t_i e^{-2\delta \sum_{i=1}^n t_i}} \right\}$, 经过化简,得

$$L(\theta, \delta) = 2n \ln \theta - 2n \ln \delta + 2n \left(\frac{\delta}{\theta} - 1 \right) \tag{5}$$

关于等式两边对参数 θ 同时求后验期望,得 $E(L(\theta, \delta)) = 2nE(\ln \theta | t) - 2n \ln \delta + 2n \left(\delta E\left(\frac{1}{\theta} | t\right) - 1 \right)$ \tag{6}

当(6)式达到最小时,可得 θ 在熵损失函数下唯一的 Bayes 估计为 $\hat{\theta}_{BE} = \left[E\left(\frac{1}{\theta} | t\right) \right]^{-1}$ \tag{7}

本文选取 θ 的先验分布为指数分布,其密度函数为 $\pi(\theta) = \lambda e^{-\lambda \theta}, \theta > 0$ \tag{8}

这里 $\lambda > 0$ 为超参数。由 Bayes 公式,可得 θ 的后验密度函数为

$$\pi(\theta | t) = \frac{4^n \theta^{2n} \lambda \prod_{i=1}^n t_i e^{-\left(2 \sum_{i=1}^n t_i + \lambda\right) \theta}}{4^n \lambda \prod_{i=1}^n t_i \int_0^{\infty} \theta^{2n} e^{-\left(2 \sum_{i=1}^n t_i + \lambda\right) \theta} d\theta} = \frac{\left(2 \sum_{i=1}^n t_i + \lambda\right)^{2n+1} \theta^{2n} \cdot e^{-\left(2 \sum_{i=1}^n t_i + \lambda\right) \theta}}{(2n)!}, \theta > 0 \tag{9}$$

定理 1 对给定的先验分布(8),在熵损失函数(5)下,艾拉姆咖参数 θ 的 Bayes 估计为 $\hat{\theta}_{BE} = \frac{2n}{2 \sum_{i=1}^n t_i + \lambda}$, 并

且该估计是唯一的。

证明 $E\left(\frac{1}{\theta} | t\right) = \frac{\left(2 \sum_{i=1}^n t_i + \lambda\right)^{2n+1}}{(2n)!} \int_0^{\infty} \theta^{2n-1} \cdot e^{-\left(2 \sum_{i=1}^n t_i + \lambda\right) \theta} d\theta = \frac{2 \sum_{i=1}^n t_i + \lambda}{2n}$

再由(7)式,得熵损失函数下 θ 的 Bayes 估计为 $\hat{\theta}_{BE} = \left[E\left(\frac{1}{\theta} | t\right) \right]^{-1} = \frac{2n}{2 \sum_{i=1}^n t_i + \lambda}$ 。证毕

2.2 Linex 损失函数下 θ 的 Bayes 估计

Linex 损失函数的表达式为 $L(\hat{\theta} - \theta) = b(e^{c(\hat{\theta} - \theta)} - c(\hat{\theta} - \theta) - 1), c \neq 0, b > 0$ 。关于等式两边对参数 θ 同时求后验期望,得

$$E(L(\hat{\theta} - \theta)) = b(e^{c\hat{\theta}} E(e^{-c\theta} | t) + cE(\theta | t) - c\hat{\theta} - 1) \tag{10}$$

当(10)式达到最小时,可得 θ 在 Linex 损失函数下的 Bayes 估计 $\hat{\theta}_{BL}$: $\hat{\theta}_{BL} = -c^{-1} \ln E(e^{-c\theta} | t)$ \tag{11}

再由文献[5]知这样的估计是唯一的。

定理 2 对于给定的先验分布(8),在 Linex 损失函数下,艾拉姆咖分布参数 θ 的 Bayes 估计为 $\hat{\theta}_{BL} = -(2n +$

$1)c^{-1} \ln \frac{2 \sum_{i=1}^n t_i + \lambda}{2 \sum_{i=1}^n t_i + \lambda + c}$, 并且该估计是唯一的。

证明 $E(e^{-c\theta} | t) = \frac{\left(2 \sum_{i=1}^n t_i + \lambda\right)^{2n+1}}{(2n)!} \int_0^{\infty} \theta^{2n} \cdot e^{-\left(2 \sum_{i=1}^n t_i + \lambda + c\right) \theta} d\theta = \left(\frac{2 \sum_{i=1}^n t_i + \lambda}{2 \sum_{i=1}^n t_i + \lambda + c} \right)^{2n+1}$ 。由(11)式得 $\hat{\theta}_{BL} =$

$$-c^{-1} \ln E(e^{-c\theta} | t) = -(2n + 1)c^{-1} \ln \frac{2 \sum_{i=1}^n t_i + \lambda}{2 \sum_{i=1}^n t_i + \lambda + c}。 \tag{证毕}$$

2.3 二次损失函数下 θ 的 Bayes 估计

取二次损失函数为 $L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2 / \theta^2$ \tag{12}

损失函数(12)消除了量纲,避免了数据单位的影响,提高了不同数据分析的可靠程度。在上述二次损失函数下,对于任何先验分布,可得 θ 的 Bayes 估计。

引理 1^[6] 在给定的先验分布和二次损失函数下, θ 的唯一的 Bayes 估计为 $\hat{\theta}_{BQ} = \frac{E(\theta^{-1} | t)}{E(\theta^{-2} | t)}$ 。

定理 3 对于给定的先验分布(8),在二次损失函数下,艾拉姆咖分布参数 θ 的 Bayes 估计为 $\hat{\theta}_{BQ} = \frac{2n-1}{2 \sum_{i=1}^n t_i + \lambda}$ 。

证明 由引理 1 得
$$\hat{\theta}_{BQ} = \frac{E(\theta^{-1} | t)}{E(\theta^{-2} | t)} = \frac{2n-1}{2 \sum_{i=1}^n t_i + \lambda} \quad \text{证毕}$$

2.4 平方损失函数下 θ 的 Bayes 估计

取平方损失函数为 $L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$, 由于在平方损失函数下, θ 的 Bayes 估计为其后验分布的均值,即

$$\hat{\theta}_{BS} = E(\theta | t) \quad (13)$$

定理 4 对于给定的先验分布(8),在平方损失函数下,艾拉姆咖分布参数 θ 的 Bayes 估计为 $\hat{\theta}_{BS} = \frac{2n+1}{2 \sum_{i=1}^n t_i + \lambda}$, 并且该估计是唯一的。

证明 由(13)式
$$\hat{\theta}_{BS} = E(\theta | t) = \frac{(2 \sum_{i=1}^n t_i + \lambda)^{2n+1}}{(2n)!} \int_0^\infty \theta^{2n+1} \cdot e^{-(2 \sum_{i=1}^n t_i + \lambda)\theta} d\theta = \frac{2n+1}{2 \sum_{i=1}^n t_i + \lambda} \quad \text{证毕}$$

2.5 平衡损失函数下 θ 的 Bayes 估计

取平衡损失函数为 $U(\hat{\theta}, \theta) = \omega (\hat{\theta} - \hat{\theta}_{BQ})^2 + (1-\omega) (\hat{\theta} - \theta)^2$, 其中, $\hat{\theta}_{BQ}$ 为二次损失函数下 θ 的 Bayes 估计, ω 为事先给定的权重。由文献[6]知,在平衡损失函数下, θ 的唯一 Bayes 估计为

$$\hat{\theta}_B = \omega \hat{\theta}_{BQ} + (1-\omega) E(\theta | t) \quad (14)$$

定理 5 对于给定的先验分布(8),在平衡损失函数下艾拉姆咖分布参数 θ 的 Bayes 估计为 $\hat{\theta}_B = \frac{2n+1-2\omega}{2 \sum_{i=1}^n t_i + \lambda}$ 。

证明 由(14)式
$$\hat{\theta}_B = \omega \hat{\theta}_{BQ} + (1-\omega) E(\theta | t) = \omega \hat{\theta}_{BQ} + (1-\omega) \hat{\theta}_{BS} = \frac{2n+1-2\omega}{2 \sum_{i=1}^n t_i + \lambda} \quad \text{证毕}$$

3 超参数 λ 的估计

对艾拉姆咖分布的密度函数,见(2)式,若取(8)式作为 θ 的先验分布,则它的边缘分布为 $f_G(t) = \int_0^{+\infty} f(t, \theta) \pi(\theta) d\theta = \frac{8\lambda t}{(2t+\lambda)^3}$ 。以 $f_G(t_i)$ 代替(3)式中的 $f(t_i, \theta)$, 则似然函数(3)式可变为 $L = \prod_{i=1}^n \frac{8t_i \lambda}{(2t_i + \lambda)^3}$, $\frac{\partial \ln L}{\partial \lambda} =$

$$\frac{n}{\lambda} - \sum_{i=1}^n \frac{3}{2t_i + \lambda} \text{。由 } \frac{\partial \ln L}{\partial \lambda} = 0 \text{ 得到 } \frac{n}{\lambda} = \sum_{i=1}^n \frac{3}{2t_i + \lambda} \text{。}$$

令 $g_1(\lambda) = \frac{n}{\lambda}$, $g_2(\lambda) = \sum_{i=1}^n \frac{3}{2t_i + \lambda}$, 显然 $g_1(\lambda) > 0, g_2(\lambda) > 0$ 。当 $\lambda \rightarrow 0^+$ 时, $g_1(\lambda) > g_2(\lambda)$, $g_1'(\lambda) = -\frac{n}{\lambda^2} < 0, g_1''(\lambda) = \frac{2n}{\lambda^3} > 0$ 。所以 $g_1(\lambda)$ 在 $(0, +\infty)$ 上为严格单调递减的下凸函数。 $g_2'(\lambda) = -\sum_{i=1}^n \frac{3}{(2t_i + \lambda)^2} < 0, g_2''(\lambda) =$

$\sum_{i=1}^n \frac{6}{(2t_i + \lambda)^3} > 0$, 所以 $g_2(\lambda)$ 在 $(0, +\infty)$ 上为严格单调递减的下凸函数。并且 $\lim_{\lambda \rightarrow +\infty} \frac{g_1(\lambda)}{g_2(\lambda)} = \lim_{\lambda \rightarrow +\infty} \frac{\frac{n}{\lambda}}{\sum_{i=1}^n \frac{3}{2t_i + \lambda}} =$

$$\frac{1}{3} < 1, \text{ 故方程 } \frac{n}{\lambda} = \sum_{i=1}^n \frac{3}{2t_i + \lambda}. \text{ 在 } (0, +\infty) \text{ 内有唯一解, 迭代公式为 } \lambda^{(p+1)} = \frac{n}{\sum_{i=1}^n \frac{3}{2t_i + \lambda^{(p)}}} \quad (15)$$

记(15)式得到的最终迭代值为 $\hat{\lambda}$, 这样就得到了超参数 λ 的估计值。

4 Bayes 估计的容许性

引理 2^[4] 在给定的 Bayes 决策问题中, 假如对给定的先验分布 $\pi(\theta)$, θ 的 Bayes 估计是唯一的, 则它是容许的。

定理 6 对于艾拉姆咖分布, 取(8)式为 θ 的先验分布, 在熵损失、Linex 损失、二次损失、平方损失和平衡损失函数下, θ 的 Bayes 估计都是容许的。

证明 由前面的定理可知在熵损失、Linex 损失、二次损失、平方损失和平衡损失函数下, θ 的 Bayes 估计都是唯一的。再根据引理 2 得到 θ 的 Bayes 估计都是容许的。

5 随机模拟及实例分析

类似文献[7], 使用 Matlab 软件通过 Monte-Carlo 模拟产生两组服从艾拉姆咖分布(1)的随机样本, 参数真值 $\theta = \frac{1}{3}$, 样本容量分别为 30 和 50。通过计算得到参数 θ 在不同损失函数下的估计值, 重复上述步骤 1 000 次, 得到参数估计的均值和均方误差, 相关结果见表 1。

表 1 参数估计的均值和均方误差

n	$\hat{\theta}_{MLE}$	$MSE(\hat{\theta}_{MLE})$	$\hat{\theta}_{BE}$	$MSE(\hat{\theta}_{BE})$	$\hat{\theta}_{BL}$	$(c=0.1 \quad -0.2)$	$MSE(\hat{\theta}_{BL})$	$(c=0.1 \quad -0.2)$
30	0.335 4	8.104e-4	0.331 0	7.696e-4	0.336 4	0.336 7	7.986e-4	8.030e-4
50	0.334 0	4.513e-4	0.331 3	4.409e-4	0.334 6	0.334 7	4.469e-4	4.482e-4
n	$\hat{\theta}_{BQ}$	$MSE(\hat{\theta}_{BQ})$	$\hat{\theta}_{BS}$	$MSE(\hat{\theta}_{BS})$	$\hat{\theta}_B$	$(\omega=0.3 \quad 0.6)$	$MSE(\hat{\theta}_B)$	$(\omega=0.3 \quad 0.6)$
30	0.325 5	8.004e-4	0.336 5	8.000e-4	0.333 2	0.329 9	7.744e-4	7.708e-4
50	0.328 0	4.506e-4	0.334 6	4.473e-4	0.332 6	0.330 7	4.408e-4	4.423e-4

由表 1 可知, 在 Linex 损失下, 当参数 $c > 0$ 和 $c < 0$ 时, θ 的 Bayes 估计差别不大; 从均方误差来看, θ 的贝叶斯估计要比极大似然估计的均方误差小; 从参数 θ 的估计与真值的接近程度及均方误差两方面来看, 在熵损失和平衡损失函数下 θ 的 Bayes 估计是较好的。

再看下面的实例。

例 1^[1] 在某型坦克维修过程中, 经过 47 次观察得到基层 I 级预防性维修二级保养时间的观测值如下:

0.80, 1.00, 1.00, 1.41, 1.50, 1.50, 1.50, 2.00, 2.00, 2.00, 2.00, 2.50, 2.50, 2.75, 3.20, 3.30, 3.70, 3.80, 3.80, 4.00, 4.00, 4.00, 4.00, 4.00, 4.00, 4.00, 4.10, 5.00, 5.00, 5.50, 5.50, 5.50, 6.00, 6.50, 7.00, 7.16, 7.75, 8.00, 8.00, 9.50, 9.73, 10.00, 11.40, 12.00, 12.00, 14.00, 15.21, 15.50。对以上数据进行统计分析, 相关结论如表 2。

表 2 参数的估计

$\hat{\theta}_{MLE}$	$\hat{\theta}_{BE}$	$\hat{\theta}_{BL}(c=-0.5)$	$\hat{\theta}_{BL}(c=0.7)$	$\hat{\theta}_{BQ}$	$\hat{\theta}_{BS}$	$\hat{\theta}_B(\omega=0.3)$	$\hat{\theta}_B(\omega=0.8)$
0.183 2	0.181 8	0.183 8	0.183 6	0.179 8	0.183 7	0.182 5	0.180 6

由表 2 可知, 几类估计值相差不大, 最大的为 0.183 8, 最小的为 0.179 8, 极差为 0.004。

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.

Lv H Q, Gao L H, Chen C L. Зрланга distribution and its application in supportability data analysis[J]. Journal of Armored Force Engineering Institute, 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] 顾蓓青,王蓉华,徐晓岭. 艾拉姆咖分布的统计分析[C]//2011年全国机械行业可靠性技术学会暨第四届可靠性工程分会第三次全体委员大会论文集, 2011: 65-67.
Gu B Q, Wang R H, Xu X L. Statistical analysis of Эрланга distribution[C]//The national machinery industry in 2011 to reliability technology and the fourth reliability engineering branch of the third members of the conference, 2011: 65-67.
- [4] 熊常伟,张德然,张怡. 熵损失函数下几何分布可靠度的 Bayes 估计[J]. 数理统计与管理, 2008, 27(1): 82-86.
Xiong C W, Zhang D R, Zhang Y. Bayesian estimation of geometric distribution parameter under entropy loss function[J]. Application of Statistics and Management, 2008, 27(1): 82-86.
- [5] 方爱秋,朱宁,李兵,等. Linex 损失函数下正态总体位置参数的估计[J]. 河南师范大学学报:自然科学版, 2008, 36(5): 5-8.
Fang A Q, Zhu N, Li B. Estimation of locate parameter of normal distribution under linex loss function[J]. Journal of Henan Normal University: Natural Science, 2008, 36(5): 5-8.
- [6] 孙玉莹,王德辉. 不同损失函数下偏正态分布的 Bayes 估计[J]. 吉林大学学报:理学版, 2012, 50(4): 638-646.
Sun Y Y, Wang D H. Bayesian estimation for the half-normal distribution under different loss function[J]. Journal of Jilin University: Science Edition, 2012, 50(4): 638-646.
- [7] 李泽华,吴小腊,刘万荣. 变系数 EV 模型系数参数核估计的改进估计[J]. 重庆师范大学学报:自然科学版, 2010, 27(1): 47-52.
Li Z H, Wu X L, Liu W R. The improvement estimation of kernel smoothing estimation in varying-coefficients EV models[J]. Journal of Chongqing Normal University: Natural Science, 2010, 27(1): 47-52.
- [8] 屈云利,朱永忠. S-SMART 最大熵法及在小样本中的应用[J]. 重庆理工大学学报:自然科学版, 2012, 26(6): 107-110.
Qu Y L, Zhu Y Z. Application of S-SMART maximum entropy method in small sample[J]. Journal of Chongqing University of Technology: Natural Science, 2012, 26(6): 107-110.
- [9] 龙兵. Rayleigh 型元件中贮备系统可靠性的近似置信下限[J]. 火力与指挥控制, 2012, 37(11): 76-80.
Long B. Lower confidence limit for reliability of a standby system with Rayleigh distribution components[J]. Fire Control & Command Control, 2012, 37(11): 76-80.

Bayesian Estimation of Эрланга Distribution Parameter under Different Loss Function—All Samples

LONG Bing

(Department of Mathematics and Physics, Jingchu University of Technology, Jingmen Hubei 448000, China)

Abstract: Эрланга distribution plays an important role in equipment repair theory. Firstly, in this paper, maximum likelihood estimation of the parameter is given under full sample on this kind of distribution. Because the performance of parameter Bayesian estimation is associated with loss function, we give the Bayesian estimation of the parameter, such as $2n / \left(2 \sum_{i=1}^n t_i + \lambda \right)$ etc under Entropy, Linex, Quadratic, Square and Balanced loss functions, and prove that the estimates we have derived are admissible. Compares mean square errors and deviation from the true values about several estimations. Bayesian estimation of θ is better under entropy loss and balanced loss function. Finally, the values about several estimations are calculated and analyzed through an example; little difference is shown in this estimation.

Key words: Эрланга distribution; prior distribution; loss function; Bayesian estimation

(责任编辑 游中胜)