

熵损失下失效率的 E-Bayes 估计及其性质*

许道军, 李国望, 沈 浮

(解放军陆军军官学院 基础部数学教研室, 合肥 230031)

摘要:失效率的多层 Bayes 估计需要用到复杂的积分, 计算比较困难。本文提出了一种新的估计方法——E-Bayes 估计法。它是在 Bayes 估计的基础上, 先给出超参数的先验分布, 进而对失效率进行估计。文章分别给出了在特定超参数先验分布的条件下, 失效率的 E-Bayes 估计和多层 Bayes 估计的计算公式。结果表明, 失效率的 E-Bayes 估计避免了多层 Bayes 估计复杂的积分计算, 形式上更加简洁, 便于计算。并通过实例证明, 对于同一组实验数据, 失效率的 E-Bayes 估计和多层 Bayes 估计的数值计算结果十分接近。这说明, 失效率的 E-Bayes 估计不仅具有多层 Bayes 估计的稳健性, 而且具有多层 Bayes 估计的精确性, 从而表明本文提出的失效率的 E-Bayes 估计法是可行的, 且相比于失效率的多层 Bayes 估计更加简洁, 更便于应用。

关键词:失效率; 熵损失; 指数分布; E-Bayes 估计; 多层 Bayes 估计

中图分类号: O213.2

文献标志码: A

文章编号: 1672-6693(2013)03-0077-04

关于参数估计, 近年来用 Bayes 方法取得了一些进展, 特别是自 Lindley 和 Smith^[1] 提出多层先验分布的想法以来, 多层 Bayes 方法在可靠性参数估计上取得了一些进展^[2-10]。但用多层 Bayes 估计参数得到的结果一般都要涉及积分计算, 使其在实际应用上还不太方便, 这在一定程度上制约了多层 Bayes 方法的应用。那么在各种方法计算结果比较接近时, 估计方法的易算性就显得尤为重要。

本文在给定失效率的先验分布下, 提出了一种新的参数估计的方法——E-Bayes 估计 (Expected Bayesian estimation) 法。给出了熵损失下, 基于有失效数据的产品失效率的 E-Bayes 估计的定义, 表达式及其性质。并通过数值例子表明, 利用 E-Bayes 估计法进行参数估计, 不仅计算结果与多层 Bayes 估计的结果十分接近, 而且不会涉及到积分计算。

设某产品的寿命 X 服从指数分布, 其密度函数为 $f(t) = \lambda \exp(-\lambda t)$ (1)

其中 $t > 0, 0 < \lambda < +\infty, \lambda$ 为指数分布(1)式的失效率。

定义 1 设随机变量 X 服从密度函数为 $f(x, \theta)$ 的分布, 其中 θ 为参数。若 δ 是 θ 的判决空间的一个估计, 则熵损失函数为似然比对数的数学期望, 即

$$L(\theta, \delta) = E_{\theta} \left[\ln \frac{f(X, \theta)}{f(X, \delta)} \right].$$

设 X_1, X_2, \dots, X_n 是来自总体(1)式的简单随机样本, 则指数分布的熵损失为

$$L(\lambda, \delta) = E_{\lambda} \left[\ln \frac{f(X_1, X_2, \dots, X_n, \lambda)}{f(X_1, X_2, \dots, X_n, \delta)} \right] =$$

$$E_{\lambda} \left[\ln \left(\frac{\lambda}{\delta} \right)^n + (\delta - \lambda) \sum_{i=1}^n X_i \right] = n \left(\ln \frac{\lambda}{\delta} + \frac{\delta}{\lambda} - 1 \right)$$

易知, 此损失函数关于 δ 是严格凸的。

引理^[4] 在熵损失下, 若 δ 是 θ 的判决空间的一个估计, 则对于任意的先验分布 $\pi(\theta), \theta$ 的 Bayes 估计为 $\hat{\delta}_B = [E(\theta^{-1} | X)]^{-1}$, 并且是唯一的。

在(1)式中, 取失效率 λ 的先验分布为 Gamma 分布(λ 的共轭先验分布)—— $Gamma(a, b)$, 其密度函数为

$$\pi(\lambda | a, b) = \frac{b^a \lambda^{a-1} \exp(-b\lambda)}{\Gamma(a)} \quad (2)$$

其中 $0 < \lambda < +\infty, a > 0, b > 0, \Gamma(a) = \int_0^{+\infty} t^{a-1} \exp(-t) dt$

为 Gamma 函数, a, b 为超参数。根据文献[2], 应选取 a, b 使 $\pi(\lambda | a, b)$ 为 λ 的减函数, 所以取 $0 < a < 1, b > 0$ 。

另外, 考虑到尾部越细的先验分布会使 Bayes 估计的稳健性越差, 因此在 $0 < a < 1$ 时 b 不易过大, 设 b 的上界为 $c(c > 0$ 为常数), 这样可以确定超参数 a, b 的取值范围为 $D = \{(a, b) | 0 < a < 1, 0 < b < c\}$ 。

* 收稿日期: 2012-06-05 修回日期: 2012-08-24 网络出版时间: 2013-05-20 18:04

资助项目: 安徽省高校优秀青年人才基金项目(No. 2011SQRL172)

作者简介: 许道军, 男, 讲师, 硕士, 研究方向为可靠性、保险精算, E-mail: candy_xdj@163.com

网络出版地址: http://www.cnki.net/kcms/detail/50.1165.N.20130520.1804.201303.77_014.html

1 失效率的 E-Bayes 估计及其性质

定义 2 设 $D = \{(a, b) | 0 < a < 1, 0 < b < c\}$, $\pi(a, b)$ 是 a, b 在区域 D 上的联合密度函数, $\hat{\lambda}(a, b)$ 为 λ 的 Bayes 估计(用超参数 a, b 表示), 称

$$\hat{\lambda}_{EB} = \iint_D \hat{\lambda}(a, b) \pi(a, b) da db$$

为 λ 的 E-Bayes 估计。

由以上定义知 λ 的 E-Bayes 估计就是 λ 的 Bayes 估计 $\hat{\lambda}(a, b)$ 对超参数 a, b 的数学期望 $E[\hat{\lambda}(a, b)]$ 。

对某产品进行 m 次定时截尾试验, 截尾时间为 $t_i (t_1 < t_2 < \dots < t_m)$, 相应试验样品数为 $n_i (i = 1, 2, \dots, m)$, 若试验的结果是 n_i 个样品中有 $r_i (0 \leq r_i \leq n_i)$ 个失效, 则称 (n_i, t_i, r_i) 为 m 次定时截尾试验的结果。为了说明超参数先验分布的变化对 λ 的 E-Bayes 估计的影响, 以下将在超参数的 3 种不同的先验分布下给出超参数 a, b 的 3 个先验分布, 并在此基础上给出 λ 的 E-Bayes 估计。

定理 1 对寿命服从指数分布(1)的产品进行 m 次定时截尾实验, 其结果为 $(n_i, t_i, r_i), i = 1, 2, \dots, m$, 在有失效数据情况下, 记 $N = \sum_{i=1}^m n_i t_i, r = \sum_{i=1}^m r_i \geq 1$, 若 λ 的先验密度函数 $\pi(\lambda | a, b)$ 由(2)式给出, 则有

1) 在熵损失下, λ 的 Bayes 估计为 $\hat{\lambda}(a, b) = \frac{a+r-1}{b+N}$;

2) 若超参数 a, b 在 D 上的先验密度函数 $\pi(a, b)$ 分别为

$$\pi_1(a, b) = \frac{2(b-1)}{c(c-2)}, (a, b) \in D \quad (3)$$

$$\pi_2(a, b) = \frac{1}{c}, (a, b) \in D \quad (4)$$

$$\pi_3(a, b) = \frac{2b}{c^2}, (a, b) \in D \quad (5)$$

则 λ 的 E-Bayes 估计分别为

$$\hat{\lambda}_{EB1} = \frac{2}{c(c-2)} \left(r - \frac{1}{2}\right) \left[c - (N+1) \ln \frac{c+N}{N} \right]$$

$$\hat{\lambda}_{EB2} = \frac{1}{c} \left(r - \frac{1}{2}\right) \ln \frac{c+N}{N}$$

$$\hat{\lambda}_{EB3} = \frac{2}{c^2} \left(r - \frac{1}{2}\right) \left[c - N \ln \frac{c+N}{N} \right]$$

证明 1) 根据文献[3], λ 的似然函数

$$L(\lambda | r) = \prod_{i=1}^m \frac{(n_i t_i)^{r_i}}{r_i!} \lambda^r \exp(-N\lambda) \propto \lambda^r \exp(-N\lambda)$$

若 λ 的先验密度函数 $\pi(\lambda | a, b)$ 由(2)式给出, 根据 Bayes 定理, λ 的后验密度函数为

$$h(\lambda | N) \propto \pi(\lambda | a, b) L(\lambda | r) =$$

$$\lambda^{a+r-1} \exp\{-(b+N)\lambda\}$$

即 $\lambda | N \sim \text{Gamma}(a+r, b+N)$, 因此在熵损失下 λ 的 Bayes 估计为

$$\hat{\lambda}(a, b) = \left[\int_0^{+\infty} \frac{1}{\lambda} h(\lambda | N) d\lambda \right]^{-1} = \frac{\Gamma(a+r)}{(b+N)\Gamma(a+r-1)} = \frac{a+r-1}{b+N}$$

2) 若超参数 a, b 在 D 上的先验密度函数 $\pi(a, b)$ 由(3)式给出, 根据定义 2, λ 的 E-Bayes 估计为

$$\begin{aligned} \hat{\lambda}_{EB1} &= \iint_D \hat{\lambda}(a, b) \pi_1(a, b) da db = \\ &= \frac{2}{c(c-2)} \int_0^1 (a+r-1) da \int_0^c \frac{b-1}{b+N} db = \\ &= \frac{2}{c(c-2)} \left(r - \frac{1}{2}\right) \left[c - (N+1) \ln \frac{c+N}{N} \right] \end{aligned}$$

同理可证, 若超参数 a, b 在 D 上的先验密度函数 $\pi(a, b)$ 由(4)、(5)式给出, 则 λ 的 E-Bayes 估计分别为

$$\hat{\lambda}_{EB2} = \frac{1}{c} \left(r - \frac{1}{2}\right) \ln \frac{c+N}{N}$$

$$\hat{\lambda}_{EB3} = \frac{2}{c^2} \left(r - \frac{1}{2}\right) \left[c - N \ln \frac{c+N}{N} \right] \quad \text{证毕}$$

定理 2 若超参数 a, b 在 D 上的先验密度函数 $\pi(a, b)$ 分别由(3)~(5)式给出, 则当 $c > 2$ 时, 有 $\hat{\lambda}_{EB2} > \hat{\lambda}_{EB3} > \hat{\lambda}_{EB1}$ 。

证明 先证 $\hat{\lambda}_{EB2} > \hat{\lambda}_{EB3}$, 由定理 1 有

$$\hat{\lambda}_{EB2} - \hat{\lambda}_{EB3} = \frac{N}{c^2} \left(r - \frac{1}{2}\right) \cdot$$

$$\left[\left(2 + \frac{c}{N}\right) \ln \left(1 + \frac{c}{N}\right) - 2 \frac{c}{N} \right] \triangleq \frac{N}{c^2} \left(r - \frac{1}{2}\right) f\left(\frac{c}{N}\right)$$

其中 $f(x) = (2+x) \ln(1+x) - 2x, x > 0$, 欲证 $\hat{\lambda}_{EB2} > \hat{\lambda}_{EB3}$, 只需证当 $x > 0$ 时, $f(x) > 0$ 。由于 $f'(x) = \ln(1+x) - \frac{x}{1+x}, f''(x) = \frac{x}{(1+x)^2}$, 故当 $x > 0$ 时, $f''(x) > 0$, 从而 $f'(x) > f'(0) = 0$, 进而 $f(x) > f(0) = 0$, 故当 $c > 0, N > 0$ 时, 有 $f\left(\frac{c}{N}\right) > 0$, 即 $\hat{\lambda}_{EB2} > \hat{\lambda}_{EB3}$ 。

再证 $\hat{\lambda}_{EB3} > \hat{\lambda}_{EB1}$, 由定理 1 可得

$$\hat{\lambda}_{EB3} - \hat{\lambda}_{EB1} = \frac{N}{c^2(c-2)} \left(r - \frac{1}{2}\right) \cdot$$

$\left[\left(2 + \frac{c}{N}\right) \ln \left(1 + \frac{c}{N}\right) - 2 \frac{c}{N} \right] \triangleq \frac{N}{c^2(c-2)} \left(r - \frac{1}{2}\right) f\left(\frac{c}{N}\right)$ 由以上推导可知, 当 $c > 2, N > 0$ 时, 有 $\hat{\lambda}_{EB3} > \hat{\lambda}_{EB1}$ 。证毕

这里需要指出的是, 定理 2 中, $c > 2$ 的要求只是为了保证(3)式有意义。

若记 $R(t) = P(X > t), t > 0$ 为产品的可靠度, 则在产品寿命服从(1)式的指数分布下, 依据失效率 λ 的 E-Bayes 估计, 得到产品可靠度的估计为 $\hat{R}_{EB}(t) = \exp(-\hat{\lambda}_{EB} t)$, 则根据定理 2, 可得如下推论。

推论 若超参数 a, b 在 D 上的先验密度函数 $\pi(a, b)$ 分别由(3)~(5)式给出,则当 $c > 2$ 时,对相同的时刻 $t > 0$,有 $\hat{R}_{EB2} < \hat{R}_{EB3} < \hat{R}_{EB1}$ 。

2 失效率的多层 Bayes 估计

为了比较 λ 的 E-Bayes 估计和 λ 的多层 Bayes 估计,以下给出 λ 的多层 Bayes 估计。若 λ 的先验密度函数 $\pi(\lambda|a, b)$ 由(2)式给出,那么超参数 a, b 如何确定呢? 文献[1]提出了多层先验分布的想法,即在先验分布中含有超参数时,可对超参数再给出一个先验分布。若超参数 a, b 的先验分布密度函数 $\pi(a, b)$ 分别由(3)~(5)式给出,则 λ 的多层先验密度函数分别为

$$\pi_1(\lambda) = \frac{2}{c(c-2)} \int_0^1 \int_0^c \frac{(b-1)b^a \lambda^{a-1} \exp(-b\lambda)}{\Gamma(a)} da db \quad (6)$$

$$\pi_2(\lambda) = \frac{1}{c} \int_0^1 \int_0^c \frac{b^a \lambda^{a-1} \exp(-b\lambda)}{\Gamma(a)} da db \quad (7)$$

$$\pi_3(\lambda) = \frac{2}{c^2} \int_0^1 \int_0^c \frac{b^{a+1} \lambda^{a-1} \exp(-b\lambda)}{\Gamma(a)} da db \quad (8)$$

定理 3 对寿命服从指数分布(1)的产品进行 m 次定时截尾试验,观察其结果为 $(n_i, t_i, r_i), i=1, 2, \dots, m$,在有失效数据情况下,若 λ 的多层先验密度函数 $\pi(\lambda)$ 分别由(6)~(8)式给出,则在熵损失下 λ 的多层 Bayes 估计分别为

$$\hat{\lambda}_{HB1} = \frac{\int_0^1 \int_0^c \frac{(b-1)b^a \Gamma(a+r)}{(b+N)^{a+r} \Gamma(a)} da db}{\int_0^1 \int_0^c \frac{(b-1)b^a \Gamma(a+r-1)}{(b+N)^{a+r-1} \Gamma(a)} da db} \quad (9)$$

$$\hat{\lambda}_{HB2} = \frac{\int_0^1 \int_0^c \frac{b^a \Gamma(a+r)}{(b+N)^{a+r} \Gamma(a)} da db}{\int_0^1 \int_0^c \frac{b^a \Gamma(a+r-1)}{(b+N)^{a+r-1} \Gamma(a)} da db} \quad (10)$$

$$\hat{\lambda}_{HB3} = \frac{\int_0^1 \int_0^c \frac{b^{a+1} \Gamma(a+r)}{(b+N)^{a+r} \Gamma(a)} da db}{\int_0^1 \int_0^c \frac{b^{a+1} \Gamma(a+r-1)}{(b+N)^{a+r-1} \Gamma(a)} da db} \quad (11)$$

证明 若 λ 的多层先验密度函数 $\pi(\lambda)$ 由(6)式给

出,根据 Bayes 定理知 λ 的多层后验密度为

$$h_1(\lambda | N) = \frac{\pi_1(\lambda)L(\lambda | r)}{\int_0^{+\infty} \pi_1(\lambda)L(\lambda | r)d\lambda} =$$

$$\frac{\int_0^1 \int_0^c \frac{(b-1)b^a}{\Gamma(a)} \lambda^{a+r-1} \exp\{-(b-N)\lambda\} da db}{\int_0^1 \int_0^c \frac{(b-1)b^a \Gamma(a+r)}{(b+N)^{a+r} \Gamma(a)} da db}$$

则在熵损失下, λ 的多层 Bayes 估计为

$$\hat{\lambda}_{HB1} = \left[\int_0^{+\infty} \frac{1}{\lambda} h_1(\lambda | N) d\lambda \right]^{-1} =$$

$$\frac{\int_0^1 \int_0^c \frac{(b-1)b^a \Gamma(a+r)}{(b+N)^{a+r} \Gamma(a)} da db}{\int_0^1 \int_0^c \frac{(b-1)b^a \Gamma(a+r-1)}{(b+N)^{a+r-1} \Gamma(a)} da db}$$

若 λ 的多层先验密度函数 $\pi(\lambda)$ 由(7)、(8)式给出,同理可得(10)、(11)式。 证毕

值得注意的是,定理 3 中的多层 Bayes 估计 $\hat{\lambda}_{HBi}$ ($i=1, 2, 3$)的结果需要用到重积分,而定理 1 中的 E-Bayes 估计 $\hat{\lambda}_{EBi}$ ($i=1, 2, 3$)的结果则无需用到积分。

3 数值算例

文献[5]给出了某型电子元件在定时截尾试验中获得的数据,具体见表 1(其中实验时间单位为:h)。

表 1 某电子元件的定时截尾实验数据

	h								
<i>i</i>	1	2	3	4	5	6	7	8	9
<i>t_i</i>	102	149	237	452	607	891	1 024	1 258	1 447
<i>n_i</i>	3	2	4	5	7	5	6	4	5
<i>r_i</i>	0	0	0	1	1	0	2	0	1

根据文献[5],该型电子元件的寿命服从指数分布。由表 1 可得 $N = \sum_{i=1}^9 n_i t_i = 30\ 927, r = \sum_{i=1}^9 r_i = 5$,根据定理 1 和定理 3,失效率 λ 的 E-Bayes 估计和多层 Bayes 估计的计算结果见表 2($c=50, 100, 500, 1\ 000, 2\ 000, 3\ 000, 4\ 000, 5\ 000, 6\ 000$)。

表 2 失效率 λ 的 E-Bayes 估计和多层 Bayes 估计 ($\times 10^{-4}$)

<i>c</i>	50	100	500	1 000	2 000	3 000	4 000	5 000	6 000	极差
$\hat{\lambda}_{EB1}$	1.453 46	1.451 89	1.439 53	1.424 40	1.395 19	1.367 28	1.340 60	1.315 05	1.290 56	0.162 90
$\hat{\lambda}_{EB2}$	1.453 86	1.452 69	1.443 40	1.432 01	1.409 93	1.388 72	1.368 34	1.348 74	1.329 85	0.124 01
$\hat{\lambda}_{EB3}$	1.453 47	1.451 91	1.439 54	1.424 42	1.395 20	1.367 30	1.340 61	1.315 06	1.290 57	0.162 90
$\hat{\lambda}_{HB1}$	1.403 22	1.412 50	1.428 71	1.426 09	1.408 24	1.386 14	1.363 23	1.340 54	1.318 49	0.110 22
$\hat{\lambda}_{HB2}$	1.398 96	1.408 19	1.426 02	1.426 01	1.413 43	1.396 48	1.378 61	1.360 85	1.343 62	0.082 40
$\hat{\lambda}_{HB3}$	1.403 07	1.412 42	1.428 70	1.426 09	1.408 25	1.386 14	1.363 23	1.340 55	1.318 50	0.110 22
极差	0.054 90	0.044 50	0.017 38	0.007 61	0.018 24	0.029 20	0.038 01	0.045 80	0.053 06	

从表 2 可以看出,对不同的 c ,失效率 λ 的 E-Bayes 估计和多层 Bayes 估计都是稳健的(最大极差

为 1.629×10^{-5}); 对相同的 $c, \hat{\lambda}_{EBi}$ 满足定理 2 的关系, 且 $\hat{\lambda}_{EBi}$ 和 $\hat{\lambda}_{HBi}$ ($i=1, 2, 3$) 十分接近(最大极差为 0.549×10^{-5}), 但相比之下, E-Bayes 估计计算简单, 更便于应用。

参考文献:

- [1] Lindley D V, Smith A F M. Bayes estimation for the linear model[J]. *Journal of the Royal Statistical Society Series B*, 1972, 34(1): 141.
- [2] 韩明. 多层先验分布的构造及其应用[J]. *运筹与管理*, 1997, 6(3): 31-40.
Han M. The structure of hierarchical prior distribution and its applications[J]. *Operations Research and Management Science*, 1997, 6(3): 31-40.
- [3] 茆诗松, 王玲玲. 可靠性统计[M]. 上海: 华东师范大学出版社, 1984.
Mao S S, Wang L L. *Statistics on reliability*[M]. Shanghai: East China Normal University Press, 1984.
- [4] 曹晋华, 程侃. 可靠性数学引论[M]. 北京: 高等教育出版社, 2006.
Cao J H, Cen N. *Mathematical introduction to reliability* [M]. Beijing: Higher Education Press, 2006.
- [5] 肖小英, 任海平. 熵损失函数下两参数 Lomax 分布形状参数的 Bayes 估计[J]. *数学的实践与认识*, 2010, 40(5): 227-230.
Xiao X Y, Ren H P. The Bayes estimation of shape parameter of Lomax distribution under entropy loss function[J]. *Journal of Mathematics in Practice and Theory*, 2010, 40(5): 227-230.
- [6] 韩明. 可靠性工程中参数的一种估计方法[J]. *中国工程科学*, 2003, 5(3): 51-56.
Han M. An new method of parameter estimation in reliability engineering[J]. *Engineering Sciences*, 2003, 5(3): 51-56.
- [7] 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. Bayes estimation of reliability of geometric distribution under entropy loss function [J]. *Journal of Applied Statistics and Management*, 2008, 27(1): 82-86.
- [8] Han M. Estimation of reliability based on zero failure data [J]. *Pure and Applied Mathematics*, 2002, 18(2): 165-169.
- [9] 韩明. 指数分布无失效数据的 Bayes 估计多层 Bayes 估计 [J]. *宁波大学学报: 理工版*, 1995, 8(2): 9-16.
Han M. Bayesian estimate multiple Bayesian estimate of exponential distribution zero [J]. *Journal of Ningbo University: Natural Science & Engineering Edition*, 1995, 8(2): 9-16.
- [10] 李俊华, 钟太勇. 熵损失函数下指数分布参数的估计 [J]. *四川理工学院学报: 自然科学版*, 2010, 23(3): 278-279.
Li J H, Zhong T Y. Estimate of exponential distribution parameter under entropy loss function [J]. *Journal of Sichuan University of Science & Engineering: Natural Science Edition*, 2010, 23(3): 278-279.

Expected Bayesian Estimation of Failure Rate and Its Character under Entropy Loss Function

XU Dao-jun, LI Guo-wang, SHEN Fu

(Dept. of Basic Courses, Army Officer Academy of P. L. A, Hefei 230031, China)

Abstract: The hierarchical Bayesian estimation of failure rate often relies on completed integral, so it's difficult to calculate. In this paper, the expected Bayesian estimation was defined, which is based on the Bayesian estimation. The failure rate is estimated after the prior distribution of super parameters was given. The formulas of the expected Bayesian estimation and the hierarchical Bayesian estimation of failure rate were given in condition of the certain super parameters, which indicates that the expected Bayesian estimation can effectively avoid the completed integral of the hierarchical Bayesian estimation, and the expected Bayesian estimation is more concise and more convenience to calculate. Finally, calculation is performed regarding to practical problem, which shows that the expected Bayesian estimation and the hierarchical Bayesian estimation are not only both steady but also approximate. As a conclusion, we get that the expected Bayesian estimation method is feasible and easier to operate.

Key words: failure rate; entropy loss function; exponential distribution; expected Bayesian estimation; hierarchical Bayesian estimation

(责任编辑 黄颖)