Statistics and Application 统计学与应用, **2024, 13(4), 1352-1364**
Published Online August 2024 in Hans. https://www.hanspub.org/journal/sa
https://doi.org/10.12677/sa.2024.134136

Hans汉斯

平衡损失下一类寿命分布的Bayes估计

陈璇璇

云南师范大学数学学院，云南 昆明

收稿日期：2024年7月20日；录用日期：2024年8月11日；发布日期：2024年8月23日

摘 要

在逐步增加II型截尾寿命数据下，本研究深入探讨了比例危险率模型的参数以及可靠性指标的贝叶斯估计与样本预测问题。首先，通过频率方法对模型参数进行了预先估计，并分析了其相关性质。随后，在平衡损失函数框架下，本文得到了可靠性指标的贝叶斯估计，同时也得出平衡损失比一般损失更加灵活的实用性结论。本文还进行了一系列数值模拟示例，其模拟结果与理论分析相一致。以上的理论研究和示例均证实了所提出的平衡损失下贝叶斯方法的实用性和有效性。

关键词

逐步增加II型截尾样本，寿命分布，平衡损失函数，Bayes估计

Bayes Estimation for a Lifetime Distribution of the Next Class of Balanced Loss

Xuanxuan Chen

School of Mathematics, Yunnan Normal University, Kunming Yunnan

Received: Jul. 20th, 2024; accepted: Aug. 11th, 2024; published: Aug. 23rd, 2024

Abstract

This study delves into the Bayesian estimation and sample prediction issues of the proportional hazard rate model's parameters and reliability indicators under progressive type-II censored lifetime data. Firstly, a preliminary estimation of the model parameters was conducted through frequency methods, and their related properties were analyzed. Subsequently, within the framework of the balanced loss function, this paper obtained the Bayesian estimation of the reliability indicators and also concluded that the balanced loss is more flexible and practical than the general loss. A series of numerical simulation examples were also conducted, and the simulation results were consistent with the theoretical analysis. The theoretical research and examples presented above all confirm the practicality and effectiveness of the proposed Bayesian method under balanced loss.

文章引用: 陈璇璇. 平衡损失下一类寿命分布的 Bayes 估计[J]. 统计学与应用, 2024, 13(4): 1352-1364.
DOI: 10.12677/sa.2024.134136

Keywords

Progressive Type-II Censoring Samples, Lifetime Distribution, Balanced Loss Function, Bayes Estimation

1. 引言

由于现代工业产品往往具有寿命时间长、试验成本高等特点，因此在实践中为了满足试验需要，人们逐步提出了截尾寿命试验，常见的截尾试验包括定数截尾和定时截尾[1]。为了使试验方案更具有实用性，研究人员又相继提出逐步增加截尾寿命试验。这是一类效率更高、成本更低的寿命试验方式，还能防止对宝贵资源不必要的破坏，具有灵活性以及高效性。所以在当今的发展进程中，这类优良的试验方案在工程应用、水文学以及生存分析等多个领域里都有广泛的应用。逐步增加截尾试验分为逐步增加 I 型截尾和逐步增加 II 型截尾两种类型。有关逐步增加截尾试验的统计分析已经吸引了很多学者的广泛讨论，如 Algarni 等[2]讨论了逆威布尔分布下逐步增加 I 型截尾样本的经典和贝叶斯估计问题；Aljohani [3]讨论了 Chen 分布下两种逐步增加 II 型截尾方案的参数估计问题。有关逐步增加截尾寿命试验的详细介绍可参考 Balakrishnan 等的著作[4]。本文在寿命模型－比例危险率分布下，利用一类广泛的损失函数－平衡损失函数，讨论该模型的可靠性指标在逐步增加 II 型截尾样本下的贝叶斯估计问题。

首先对逐步增加 II 型截尾寿命试验模式做一简要介绍：假设有 n 个相同样品同时开始试验，并且样品失效时间均可被检测到，设 $m<n$， $r_1,r_2,\cdots,r_m$ 是试验前预先固定的一组正整数，并且满足条件 $r_1+r_2+\cdots+r_m+m=n$。当第一个失效样品出现时，记失效时刻为 X_1，同时在剩下 $n-1$ 个未失效样品中随机移离 r_1 个样品；当第二个失效样品出现时，记失效时刻为 X_2，在剩下 $n-2-r_1$ 个样品中随机移离 r_2 个样品；以同样的方法继续试验直至第 m 个失效样品出现，记失效时刻为 r_m，并移离剩下的全部 $r_m=n-m-r_1-r_2-\cdots-r_{m-1}$ 个样品，试验完毕。由定义可以看到，逐步增加 II 型截尾寿命试验是一类比定时、定数截尾试验更加灵活广泛的寿命试验。当截尾策略 $r_1=r_2=\cdots=r_{m-1}=0$ 时，该试验是通常的定数截尾试验；当 $r_1=r_2=\cdots=r_m=0$ 时，相当于完全寿命试验。设受试样品寿命服从的分布函数为 $F(x)$，密度为 $f(x)$ 的寿命分布，则相应的逐步增加 II 型截尾样本 $X=(X_1,X_2,\cdots,X_m)$ 的联合概率分布为：

$$f(x)=c\prod_{i=1}^{m}f(x_i)\left[1-F(x_i)\right]^{r_m}, \tag{1}$$

其中，c 表示独立于参数的正则化常数。

在统计决策理论中，为了描述决策与风险之间的关系，人们常常引入损失函数的概念，常见的有均方损失、熵损失、对数损失等。然而，通常的损失函数仅仅只顾及参数估计的精确性，而忽略了数据对于分布函数的拟合优度。为了在决策分析中兼顾分布函数拟合及参数估计的精确性，Zellner [5]最早引入了平衡损失函数的概念，Mohammad 等[6]在 Zellner 的基础上进一步提出了一类更加广泛的平衡损失函数，其形式如下：

$$L^q_{\rho,\omega,\delta}\left(\gamma(\theta),\delta\right)=\omega q(\theta)\rho(\delta_0,\delta)+(1-\omega)q(\theta)\rho\left(\gamma(\theta),\delta\right), \tag{2}$$

其中，$\omega\in(0,1)$，$q(\theta)$ 是正的权函数，$\rho\left(\gamma(\theta),\delta\right)$ 表示任意的损失函数，其中 δ 是 $\gamma(\theta)$ 的估计值，δ_0 是

预先给出的$\gamma(\theta)$的一个估计，如似然估计、最小二乘估计等。由定义可见，平衡损失$L^{q}_{\rho,\omega,\delta}(\gamma(\theta),\delta)$是一类更加广泛的损失函数，可以用来描述各种不同场合。当权因子$\omega\to 0$时，$L^{q}_{\rho,\omega,\delta}(\gamma(\theta),\delta)$退化为通常的损失函数$q(\theta)\rho(\gamma(\theta),\delta)$。

本文考虑样本寿命X服从如下的模型，其分布函数为：

$$F(x)=1-\left[\bar{G}(x)\right]^{\theta},-\infty\le c<x<d\le\infty \tag{3}$$

相应的密度函数为$f(x)$为：

$$f(x;\theta)=g(x)\theta\left[\bar{G}(x)\right]^{\theta-1},-\infty\le c<x<d\le\infty \tag{4}$$

其中，θ是未知模型参数，$g(x)=G'(x)>0,\bar{G}(x)=1-G(x)$是任意连续函数，且满足$G(c)=0,G(d)=1$。

模型(3)被称为比例危险率模型。

由(3)、(4)可以看到，比例危险率模型是一类非常重要的广义半参数模型，在生存分析、可靠性理论及质量控制等方面都具有广泛的应用，很多常见的寿命分布都是比例危险率模型的特殊情形，例如：

1) 当$G(x)=\exp\{-x\},x>0$时，模型(3)为指数分布，其分布密度为：

$$f(x;\theta)=\theta\exp\{-\theta x\},x>0$$

2) 当$G(x)=\exp\{-x^2\},x>0$时，模型(3)为 Rayleigh 分布，其分布密度为：

$$f(x;\theta)=2x\theta\exp\{-\theta x^2\},x>0.$$

3) 当$G(x)=\beta/x,x>\beta$时，模型(3)为 Pareto 分布，其分布密度为：

$$f(x;\theta)=\frac{\beta\theta}{x^2}\left(\frac{\beta}{x}\right)^{\theta-1},x>\beta$$

因此，区别于以往对传统数据类型的研究，本文的创新之处即在平衡损失函数下，考察当数据为逐步增加 II 型截尾样本时，比例危险率模型分布参数和可靠性指标的 Bayes 估计问题，为寿命试验领域提供了较为新颖的视角。

由(3)、(4)可知，寿命分布X在时间t的生存函数和失效率函数分别为：

$$R(t)=1-F(t;\theta)\left(\bar{G}(x)\right)^{\theta},\quad H(t)=\frac{f(t;\theta)}{1-F(t;\theta)}=g(t)\theta/\bar{G}(t). \tag{5}$$

假设$X=(X_1,X_2,\cdots,X_m)$是来自于模型(3)的逐步增加 II 型截尾样本，由(1)可知，截尾样本$X=(X_1,X_2,\cdots,X_m)$的联合似然函数可以表示为：

$$f_X(x;\theta)=c\prod_{i=1}^{m}\frac{g(x_i)}{\bar{G}(x_i)}\theta^m\exp[-W\theta], \tag{6}$$

其中，$W=-\sum_{i=1}^{m}(r_i+1)\ln\bar{G}(x_i)$。

2. Bayes 估计

本节给出参数和可靠性指标的 Bayes 估计。

2.1. 先验信息及后验分布

设参数θ的先验分布为共轭伽玛分布$\mathrm{Gam}(1,a)$：

$$\pi(\theta)=a\mathrm{e}^{-a\theta},\theta>0, \tag{7}$$

其中，$a>0$ 为超参数。

由(6)和(7)可知，参数 θ 的后验分布为：

$$\pi(\theta \mid x)=\frac{(a+W)^{m+1}}{\Gamma(m+1)}\theta^{m}\mathrm{e}^{-\theta(a+W)},\theta>0, \tag{8}$$

当 $q(\theta)=1$ 时，记平衡损失函数(2)为 $L_{\rho}(\gamma(\theta),\delta)$，本文分别在 $\rho(\gamma(\theta),\delta)$ 为均方误差和 Linex 损失情形下，讨论参数及可靠性指标的贝叶斯估计。由于利用平衡损失函数进行决策分析时需要使用到相关的预估计 δ_0，下面我们首先给出两类频率估计作为 δ_0 的选择。

2.2. 频率估计

本节给出参数及可靠性指标的极大似然估计(MLE)和一致最小方差无偏估计(UMVU)。

由(6)可知，对数似然函数为：

$$\pi(\theta;x)=\ln f_X(x;\theta)\propto m\ln\theta-W\theta, \tag{9}$$

对(9)关于 θ 求导数，利用不变性原理分别可得参数 θ、可靠性指标 $R(t),H(t)$ 的 MLE 如下：

$$\hat{\theta}_M=\frac{m}{W},\hat{R}_M=\left[\bar{G}(t)\right]^{m/W},H_M=\frac{g(t)}{\bar{G}(t)}\frac{m}{W} \tag{10}$$

为了获得 UMVU 估计，下面给出一个有用的引理。

引理 1 设 $X_1,X_2,\cdots,X_m$ 是来自于分布(3)的逐步增加 II 型截尾样本，则 $T=2\theta W$ 服从自游度为 $2m$ 的卡方分布。

证明 定义 $Y_i=-\theta\ln\left[\bar{G}(x_{ij})\right]$，则 $Y_i,i=1,2,\cdots,m$ 是来自于标准指数分布的逐步增加 II 型截尾样本。考虑如下变换：

$$\begin{aligned}
&Z_1=nY_1,\\
&Z_2=(n-r_1-1)(Y_2-Y_1),\\
&\vdots\\
&Z_m=(n-r_1-\cdots-r_{m-1}-m+1)(Y_m-Y_{m-1}),
\end{aligned}$$

则 $Z_i,i=1,2,\cdots,m$ 是来自于标准指数分布的独立同分布样本。进一步，由[1]可知：

$$2\theta W=-2\theta\sum_{i=1}^{m}(r_i+1)\ln\bar{G}(x_i)=2\sum_{i=1}^{m}(r_i+1)Y_i=2\sum_{i=1}^{m}Z_i,$$

服从自由度为的 $2m$ 的卡方分布，其密度函数为：

$$f(t)=\frac{1}{2^{2m/2}\Gamma(2m/2)}t^{2m/2-1}\mathrm{e}^{-t/2},t>0.$$

结论得证。

推论 1 参数 θ 和失效率 $H(t)$ 的极大似然估计是渐进无偏，且相合的。

证明 我们分别考察 θ 和 $H(t)$ 的 MLE 的样本性质。

利用引理 1，我们有：

$$E\left[\hat{\theta}_M\right]=E\left[\frac{m}{W}\right]=2m\theta E\left[\frac{1}{2\theta W}\right]=\frac{m}{m-1}\theta,$$

和

$$E\left[\hat{\theta}_M^2\right]=E\left[\frac{m}{W}\right]^2=4m^2\theta^2E\left[\frac{1}{2\theta W}\right]^2=\frac{m^2}{(m-1)(m-2)}\theta^2,$$

从而有参数 θ 极大似然估计的期望和方差为：

$$E\left[\hat{\theta}_M\right]=\frac{m}{m-1}\theta,\ \mathrm{Var}\left[\hat{\theta}_M\right]=\frac{m^2}{(m-1)^2(m-2)}\theta^2 \tag{11}$$

类似地，参数 $H(t)$ 极大似然估计的期望和方差为：

$$E\left[\hat{H}_M\right]=\frac{m}{m-1}H(t),\ \mathrm{Var}\left[\hat{H}_M\right]=\frac{m^2}{(m-1)^2(m-2)}H^2(t). \tag{12}$$

因为

$$\lim_{m\to\infty}E\left[\hat{\theta}_M\right]=\theta,\ \lim_{m\to\infty}\mathrm{Var}\left[\hat{\theta}_M\right]=0,$$

$$\lim_{m\to\infty}E\left[\hat{H}_M\right]=H(t),\ \lim_{m\to\infty}\mathrm{Var}\left[\hat{H}_M\right]=0,$$

从而结论成立。

因为 W 是参数 θ 的充分完备统计量，利用 Rao-Blackwell-Lehmann-Scheffe 定理和引理 1，一步可以得到参数和可靠性指标的 UMUV 估计如下：

$$\hat{\theta}_U=\frac{m-1}{W},\ \hat{R}_U(t)=\left[1+\frac{\ln\bar{G}(t)}{W}\right]^{m-1},\ \hat{H}_U(t)=\frac{g(t)}{\bar{G}(t)}\frac{m-1}{W}, \tag{13}$$

推论 2 参数 θ 和可靠性指标 $H(t)$ 的一致最小方差无偏估计是相合的。

证明 由(11)和(12)可知：

$$\mathrm{Var}\left[\hat{\theta}_U\right]=\frac{1}{m-2}\theta^2,\quad \mathrm{Var}\left[\hat{H}_U\right]=\frac{1}{m-2}H^2(t).$$

由相合性定义知，统计量 $\hat{\theta}_U,\hat{H}_U$ 是相合估计。

结合引理 1 和 $Z_1,Z_1,\cdots,Z_m$ 的独立性，下面给出几个假设检验问题。

注 1. (单参数假设检验)当获得逐步增加 II 型截尾样本 $X=(X_1,X_2,\cdots,X_m)$ 时，有时需要考虑参数 θ 如下的假设检验：

(a) $H_0:\theta\le\theta_0\leftrightarrow H_1:\theta>\theta_0$，

(b) $H_0:\theta\ge\theta_0\leftrightarrow H_1:\theta<\theta_2$，

(c) $H_0:\theta=\theta_0\leftrightarrow H_1:\theta\ne\theta_0$。

由引理 1 可知，$2\theta W$ 服从自由度为 $2m$ 的卡方分布。在检验水平 $\alpha\in(0,1)$，对上述检验(a)、(b)和(c)，相应原假设 H_0 的拒绝域分别为：

(a)′ $\left\{\frac{m}{W}\ge\frac{2m\theta_0}{\chi^2_{2m}(1-\alpha)}\right\}$，

(b)′ $\left\{\frac{m}{W}\le\frac{2m\theta_0}{\chi^2_{2m}(\alpha)}\right\}$，

(c)′ $\left\{\frac{m}{W}\le\frac{2m\theta_0}{\chi^2_{2m}\left(\frac{\alpha}{2}\right)}\text{或}\frac{m}{W}\ge\frac{2m\theta_0}{\chi^2_{2m}\left(1-\frac{\alpha}{2}\right)}\right\}$。

这里，$\chi^2_{2m}(\alpha)$ 表示自由度为 $2m$ 的卡方分布的$100\alpha\%$上侧分位数。

注 2. (两样本假设检验)这里给出一个两样本假设检验问题。在实际中，我们有时常常需要比较服从同一模型的两样本参数。比如，当某产品的寿命分布服从指数分布($\bar{G}(x)=\mathrm{e}^{-x}$)时，在进行技术革新或者设备调整之后，我们常常需要比较调整前后产品的平均寿命是否得到提高。由于在指数分布下，$1/\theta$ 表示产品的平均寿命，因此，需要比较变化前参数 θ_1 和变化后参数 θ_2 的大小。如果技术革新有效，产品的寿命提高，则有 $\theta_1>\theta_2$。下面我们给出一类两样本的假设检验问题。设 $X^{(i)}=\left(X_1^i,X_2^i,\cdots,X_m^i\right),i=1,2$ 表示来自于寿命分布(3)参数为 θ_i，样本容量为 m_i，截尾策略为 $r_i=\left(r_1^{(i)},r_m^{(i)},\cdots,r_{m_i}^{(i)}\right)$ 的逐步增加 II 型截尾样本。我们考虑以下假设检验问题：

(d) $H_0:\theta_1\le\theta_2\leftrightarrow H_1:\theta_1>\theta_2$，

(e) $H_0:\theta_1\ge\theta_2\leftrightarrow H_1:\theta_1<\theta_2$，

(f) $H_0:\theta_1=\theta_2\leftrightarrow H_1:\theta_1\ne\theta_2$。

令 $W_i=-\sum_{j=1}^{m_i}\left(r_j^{(i)}+1\right)\ln\bar{G}\left(x_j^{(i)}\right),i=1,2$，由引理 1 可知，$2\theta_iW_i$ 服从自由度为 $2m_i$ 的卡方分布，进而有 $\frac{\theta_2}{\theta_1}\frac{m_1W_2}{m_2W_1}$ 服从自由度为 $2m_2$ 和 $2m_1$ 的 F 分布。进而，在检验水平 $\alpha\in(0,1)$ 下，对上述检验(e)、(d)和(f)，相应原假设 H_0 的拒绝域分别为：

(d)′ $\left\{\frac{m_1}{m_2}\frac{W_2}{W_1}\ge F_{(2m_2,2m_1)}(\alpha)\right\}$，

(e)′ $\left\{\frac{m_1}{m_2}\frac{W_2}{W_1}\le F_{(2m_2,2m_1)}(1-\alpha)\right\}$，

(e)′ $\left\{\frac{m_1}{m_2}\frac{W_2}{W_1}\ge F_{(2m_2,2m_1)}\left(\frac{\alpha}{2}\right)\text{或}\frac{m_1}{m_2}\frac{W_2}{W_1}\le F_{(2m_2,2m_1)}\left(1-\frac{\alpha}{2}\right)\right\}$。

这里，$F_{(2m_2,2m_1)}(\alpha)$ 表示自由度为 $2m_2$ 和 $2m_1$ 的 F 分布的$100\alpha\%$上侧分位数。

2.3. 平衡均方损失函数下的 Bayes 估计

当 $\rho(\gamma(\theta),\delta)=(\delta-\gamma(\theta))^2$，由(2)及风险函数最小准则，$\gamma(\theta)$ 在平衡均方损失下 $L_s(\gamma(\theta),\delta)$ 的 Bayes 估计为：

$$\hat{\delta}_{BS}=\omega\delta_0+(1-\omega)\frac{E\left[q(\theta)\gamma(\theta)\mid x\right]}{E\left[q(\theta)\mid x\right]}. \tag{14}$$

由(8)和(14)可知，在平衡均方损失 $L_S(\gamma(\theta),\delta)$ 下参数 θ 和可靠性指标 $R(t),H(t)$ 的 Bayes 估计如下：

$$\begin{aligned}\hat{\theta}_{BS}&=\omega\delta_0+(1-\omega)\frac{m+1}{a+W},\\ \hat{R}_{BS}(t)&=\omega\delta_0+(1-\omega)\left(\frac{a+W}{a+W-\ln\bar{G}(t)}\right)^{m+1},\\ \hat{H}_{BS}(t)&=\omega\delta_0+(1-\omega)\frac{g(t)}{\bar{G}(t)}\frac{m+1}{a+W}.\end{aligned} \tag{15}$$

在参数 θ 和可靠性指标 $R(t),H(t)$ 的 Bayes 估计(15)中，估计 δ_0 表示各个指标相应的预先估计值，在下述平衡 Linex 损失中亦同。

注 4. 当先验分布(7)中超参数 $a\to0$ 时，在平衡均方损失下相应的 Bayes 估计为：

$$\hat{\theta}_{BS}^{N}=\omega\delta_0+(1-\omega)\frac{m+1}{W},$$

$$\hat{R}_{BS}^{N}(t)=\omega\delta_0+(1-\omega)\left(\frac{W}{W-\ln\bar{G}(t)}\right)^{m+1},$$

$$\hat{H}_{BS}^{N}(t)=\omega\delta_0+(1-\omega)\frac{g(t)}{\bar{G}(t)}\frac{m+1}{W}.$$

直接计算可知，$\hat{\theta}_{BS}^{N}$、$\hat{R}_{BS}^{N}(t)$和$\hat{H}_{BS}^{N}(t)$是当先验取无信息先验分布$\pi(\theta)=1/\theta$时，在平衡均方损失函数下相应的 Bayes 估计。

2.4. 平衡 Linex 损失函数下的贝叶斯估计

Linex 损失函数是近年来研究较多的一类非对称损失函数，与均方损失不同，该损失函数能够对过高估计和过低估计产生的风险区别对待，从而更加客观地描述现实情况，因而受到众多学者和实际工作者的青睐，其基本形式如下：

$$\rho(\gamma(\theta),\delta)=\mathrm{e}^{c(\delta-\gamma(\theta))}-c(\delta-\gamma(\theta))-1,c\neq 0.$$

由定义可以看到，当$c<0$时，过低估计所造成的损失大于过高估计；当$c>0$时，结论正好相反；当$c\to 0$时，非对称 Linex 损失渐近趋于均方损失，且渐近对称。由于在 Linex 损失下$\gamma(\theta)$的 Bayes 估计为$\hat{\delta}'=-(1/c)\ln\left(E\left[\mathrm{e}^{-c\gamma(\theta)}\mid x\right]\right)$，根据风险最小原则，在平衡 Linex 损失函数$\gamma(\theta)$的 Bayes 估计为：

$$\hat{\delta}_{BL}=-\frac{1}{c}\ln\left\{\omega\mathrm{e}^{-c\delta_0}+(1-\omega)\mathrm{e}^{-c\hat{\delta}'}\right\}. \tag{16}$$

由(8)和(16)可得，在平衡 Linex 损失$L_L(\gamma(\theta),\delta)$下参数$\theta$和可靠性指标$R(t),H(t)$的 Bayes 估计分别为：

$$\hat{\theta}_{BL}=-\frac{1}{c}\ln\left\{\omega\mathrm{e}^{-c\delta_0}+(1-\omega)\left(\frac{a+W}{a+W+c}\right)^{m+1}\right\},$$

$$\hat{R}_{BL}(t)=-\frac{1}{c}\ln\left\{\omega\mathrm{e}^{-c\delta_0}+(1-\omega)\sum_{k=0}^{\infty}\frac{(-c)^k}{k!}\left(\frac{a+W}{a+W-k\ln(\bar{G}(t))}\right)^{m+1}\right\}, \tag{17}$$

$$\hat{H}_{BL}(t)=-\frac{1}{c}\ln\left\{\omega\mathrm{e}^{-a\delta_0}+(1-\omega)\left(\frac{a+W}{a+W+cg(t)/\bar{G}(t)}\right)^{m+1}\right\}.$$

注 5. 当先验分布(7)超参数$a\to 0$时，在平衡均方损失下相应的 Bayes 估计为：

$$\hat{\theta}_{BL}^{N}=-\frac{1}{c}\ln\left\{\omega\mathrm{e}^{-c\delta_0}+(1-\omega)\left(\frac{W}{W+c}\right)^{m+1}\right\},$$

$$\hat{R}_{BL}^{N}(t)=-\frac{1}{c}\ln\left\{\omega\mathrm{e}^{-a\delta_0}+(1-\omega)\sum_{k=0}^{\infty}\frac{(-c)^k}{k!}\left(\frac{W}{W-k\ln\bar{G}(t)}\right)^{m+1}\right\},$$

$$\hat{H}_{BL}^{N}(t)=-\frac{1}{c}\ln\left\{\omega\mathrm{e}^{-a\delta_0}+(1-\omega)\left(\frac{W}{W+cg(t)/\bar{G}(t)}\right)^{m+1}\right\}.$$

类似于平衡均方损失，$\hat{\theta}_{BL}^{N},\hat{R}_{BL}^{N}(t)$ 和 $\hat{H}_{BL}^{N}(t)$ 是当先验分布取无信息先验分布 $\pi(\theta)=1/\theta$，在平衡 Linex 损失函数下相应的 Bayes 估计。

2.5. 超参数估计

设总体 X 是来自于分布(3)的寿命分布，令 $Y=-\ln(1-G(x))$，则 Y 服从指数分布：

$$f_Y(y;\theta)=\theta\exp\{-\theta y\},y>0.$$

进一步，令 $Y_i=-\ln(1-G(x_i)),i=1,2,\cdots,m$，则 Y_i 是来自于指数分布 $f_Y(y;\theta)$ 的逐步增加 II 型截尾样本。

其条件密度函数为：

$$f_Y(y)=\int_0^{\infty}\pi(\theta)f_Y(y;\theta)\mathrm{d}\theta=\frac{a}{(a+y)^2},y>0.$$

相应地，条件分布密度为：

$$F_Y(y)=\int_0^{y}f_Y(t)\mathrm{d}t=1-\frac{a}{a+y},y>0.$$

所以，逐步增加 II 型截尾样本 $Y_1,Y_2,\cdots,Y_m$ 的似然函数为：

$$f_{(Y_1,\cdots,Y_m)}(y)\propto\prod_{i=1}^{m}f_Y(y_i)[1-F_Y(y)]=\prod_{i=1}^{m}\frac{a}{(a+y_i)^2}\left[\frac{a}{a+y_i}\right]^{r_i}n=a^{\sum_{i=1}^{m}(r_i+1)}\prod_{i=1}^{m}(a+y_i)^{-(r_i+2)},$$

从而对数似然函数为：

$$L(a)=\ln f_{(Y_1,\cdots,Y_m)}(y)\propto\sum_{i=1}^{m}(r_i+1)\ln a-\sum_{i=1}^{m}(r_i+2)\ln(a+y_i).$$

关于对数似然函数取导数可得：

$$\frac{\partial}{\partial a}L(a)=\frac{1}{a}\sum_{i=1}^{m}(r_i+1)-\sum_{i=1}^{m}\frac{r_i+2}{a+y_i}.$$

由 ML-II 法可知，如果超参数 a 的似然估计存在，只需证明对数似然函数有唯一的根即可。

令

$$g_1(a)=\frac{1}{a}\sum_{i=1}^{m}(r_i+1),\quad g_2(a)=\sum_{i=1}^{m}\frac{r_i+2}{a+y_i}.$$

因为

$$\lim_{a\to0}g_1(a)=\infty,\lim_{a\to\infty}g_1(a)=0,$$

$$\frac{\partial}{\partial a}g_1(a)=-\frac{1}{a^2}\sum_{i=1}^{m}(r_i+1)<0,$$

$$\frac{\partial^2}{\partial a^2}g_1(a)=\frac{2}{a^2}\sum_{i=1}^{m}(r_i+1)>0,$$

且

$$\lim_{a\to0}g_2(a)=\sum_{i=0}^{m}\frac{r_i+2}{y_i},\lim_{a\to\infty}g_2(a)=0,$$

$$\frac{\partial}{\partial a}g_2(a)=-\sum_{i=1}^{m}\frac{r_i+2}{(a+y_i)^2}<0,$$

$$\frac{\partial^2}{\partial a^2}g_2(a)=2\sum_{i=1}^{m}\frac{r_i+2}{(a+y_i)^3}>0,$$

由上述两式可知，函数 $g_1(a)$ 和 $g_2(a)$ 是关于 a 的连续单调递减凸函数。

又因为

$$\lim_{a\to\infty}\frac{g_2(a)}{g_1(a)}=\lim_{a\to\infty}\frac{\sum_{i=1}^{m}\frac{r_i+2}{a+y_i}}{\frac{1}{a}\sum_{i=1}^{m}(r_i+1)}=\frac{\sum_{i=1}^{m}\lim_{a\to\infty}\frac{a(r_i+2)}{a+y_i}}{\sum_{i=1}^{m}(r_i+1)}=\frac{\sum_{i=1}^{m}r_i+2}{\sum_{i=1}^{m}(r_i+1)}>1,$$

从而方程 $(\partial/\partial a)L(a)=0$ 存在唯一解，即超参数 a 的估计存在且唯一。因为由方程 $(\partial/\partial a)L(a)=0$ 不能直接给出 a 的精确数值解，这里给出下面的迭代算法来获得超参数 a 的 ML-II 估计：

$$a^{(k+1)}=\sum_{i=1}^{m}(r_i+1)\cdot\sum_{i=1}^{m}\frac{r_i+2}{a^{(k)}-\ln(1-G(x_i))},k=0,1,2,\cdots \tag{18}$$

其中，a^k 表示 a 的第 k 次迭代值，a^0 表示 a 的一个给定的初始值。

记由(18)获得的 a 的 ML-II 估计为 $\hat{a}$，将 $\hat{a}$ 带入参数 θ 和可靠性指标 $R(t),H(t)$ 的 Bayes 估计，进而获得相关估计的经验 Bayes 估计。

2.6. Bayes 最大后验密度区间估计

这里给出参数 θ 和可靠性指标 $R(t),H(t)$ 的最大后验密度区间估计(HPD)。

给定水平值 $a\in(0,1)$，设 $[\theta_L,\theta_U]$ 是参数 θ 的 $100(1-\alpha)\%$ HPD 区间估计，则 θ_L,θ_U 满足下列条件：

$$\int_{\theta_L}^{\theta_U}\pi(\theta\,|\,x)=1-\alpha,\pi(\theta_L\,|\,x)=\pi(\theta_U\,|\,x).$$

由(8)可知：

$$\Gamma^*(m;q_1,q_2)=(1-\alpha)\Gamma(m+1),\left(\frac{\theta_L}{\theta_U}\right)^m=\exp\{q_1-q_2\}. \tag{19}$$

其中，$q_1=(a+W)\theta_L,q_2=(a+W)\theta_U,\Gamma^*(v;t_1,t_2)=\int_{t_1}^{t_2}t^v\mathrm{e}^{-t}\mathrm{d}t$。

为了得到 $R(t),H(t)$ 的 HPD 区间估计，由(8)可知，可靠性指标 $R(t),H(t)$ 的后验密度函数分别为：

$$\pi(R\,|\,x)=\frac{(a+W)^{m+1}}{\Gamma(m+1)}\frac{1}{R}\left(\frac{\ln R}{\ln\bar{G}(t)}\right)^m\frac{1}{-\ln\bar{G}(t)}\mathrm{e}^{\frac{\ln R}{\ln\bar{G}(t)}(a+W)},0<R<1,$$

$$\pi(H\,|\,x)=\frac{(a+W)^{m+1}}{\Gamma(m+1)}\left(\frac{\bar{G}(t)}{g(t)}\right)^{m+1}H^m\mathrm{e}^{-\frac{\bar{G}(t)}{g(t)}H(a+W)},H>0.$$

类似于 HPD 区间估计 $[\theta_L,\theta_U]$，可得 $R(t)$ 的 HPD 区间估计 R_L,R_U 满足：

$$\Gamma^*(m;e_1,e_2)=(1-\alpha)\Gamma(m+1),\frac{R_U}{R_L}\left(\frac{e_2}{e_1}\right)^m=\exp\{e_1-e_2\}. \tag{20}$$

其中，$e_1=(a+W)\dfrac{\ln R_U}{\ln\bar G(t)}, e_2=(a+W)\dfrac{\ln R_L}{\ln\bar G(t)}$。

进一步，$H(t)$ 的 HPD 区间估计 $[H_L,H_U]$ 满足：

$$\Gamma^*(m;d_1,d_2)=(1-\alpha)\Gamma(m+1),\left(\frac{d_1}{d_2}\right)^m=\exp\{d_1-d_2\}. \tag{21}$$

其中，$d_1=\dfrac{\bar G(t)}{g(t)}H_L(a+W), d_2=\dfrac{\bar G(t)}{g(t)}H_U(a+W)$。

3. Bayes 预测

记 $m<s\le n, Z=X_s$ 表示第 s 个逐步增加 II 型截尾样本，基于历史样本 $X=(X_1,\cdots,X_m)$ 的贝叶斯预测密度为：

$$f(z\mid x)=\int_0^\infty f(z\mid x_m,\theta)\pi(\theta\mid x)\mathrm{d}\theta,$$

其中，$f(z\mid x_m,\theta)=\dfrac{[v(z)-v(x_m)]^{s-m-1}}{\Gamma(s-m)}\dfrac{f(z;\theta)}{1-F(x_m;\theta)}, z>x_m$，这里 $v(\cdot)=-\ln[1-F(\cdot\mid\theta)]$。

由(3)、(4)可知，$f(z\mid x_m,\theta)=\dfrac{[\ln\bar G(x_m)-\ln\bar G(z)]^{s-m-1}}{\Gamma(s-m)}\dfrac{g(z)}{\bar G(z)}\left[\dfrac{\bar G(z)}{\bar G(x_m)}\right]^\theta\theta^{s-m}$。

从而 $Y=X_s$，预测密度为：

$$\begin{aligned}f(z\mid x)&=\int_0^\infty\frac{[\ln\bar G(x_m)-\ln\bar G(z)]^{s-m-1}}{\Gamma(s-m)}\frac{g(z)}{\bar G(z)}\left[\frac{\bar G(z)}{\bar G(x_m)}\right]^\theta\theta^{s-m}\cdot\frac{(a+W)^{m+1}}{\Gamma(m+1)}\theta^m\mathrm{e}^{-\theta(a+W)}\mathrm{d}\theta\\&=\frac{[\ln\bar G(x_m)-\ln\bar G(z)]^{s-m-1}}{\Gamma(s-m)}\frac{g(z)}{\bar G(z)}\frac{(a+W)^{m+1}}{\Gamma(m+1)}\cdot\int_0^\infty\theta^s\exp\left\{-\theta\left(a+W+\ln\bar G(x_m)-\ln\bar G(z)\right)\right\}\mathrm{d}\theta\\&=\frac{\Gamma(s+1)(a+W)^{m+1}}{\Gamma(s-m)\Gamma(m+1)}\frac{g(z)}{\bar G(z)}\frac{[\ln\bar G(x_m)-\ln\bar G(z)]^{s-m-1}}{[a+W+\ln\bar G(x_m)-\ln\bar G(z)]^{s+1}}.\end{aligned}$$

从而对任意 $z>x_m$，有：

$$\begin{aligned}P(Z>z\mid x)&=\int_{x_m}^z f(t\mid x)\mathrm{d}t\\&=\int_{x_m}^z\frac{\Gamma(s+1)(a+W)^{m+1}}{\Gamma(s-m)\Gamma(m+1)}\frac{[\ln\bar G(x_m)-\ln\bar G(t)]^{s-m-1}}{[a+W+\ln\bar G(x_m)-\ln\bar G(t)]^{s+1}}\mathrm{d}[\ln\bar G(x_m)-\ln\bar G(t)]\\&=\mathbb{F}\left(\frac{\ln\bar G(x_m)-\ln\bar G(z)}{a+W};s-m,m+1\right),\end{aligned}$$

其中，$\mathbb{F}(z;a,b)=\int_0^z\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\dfrac{t^{a-1}}{(1+t)^{a+b}}\mathrm{d}t$ 表示 Fisher Z 分布在点 z 的值。

记 U,L 是 $Z=X_m$ 的 $100(1-\alpha)\%$ 的 Bayes 预测上下界，满足：

$$P(L<Z<U)=1-\alpha,\alpha\in(0,1).$$

从而 X_m 的预测上下界 U,L 可由下面的式子分别得到：

$$P\left(Z>U\right)=\frac{\alpha}{2},\quad P\left(Z>L\right)=1-\frac{\alpha}{2}.$$

在实际应用中，人们往往比较关心一步预测，即 $s=m+1$ 的情形，这时有：

$$P\left(Y>y\mid x\right)=1-\left[\frac{\ln\bar{G}\left(x_m\right)-\ln\bar{G}\left(y\right)}{a+W}\right]^{m+1}.$$

从而 X_{m+1} 的一步预测上下界分别为：

$$U=\bar{G}^{-1}\left[\exp\left\{\ln\bar{G}\left(x_m\right)-\left(a+W\right)\left(1-\frac{\alpha}{2}\right)^{1/(m+1)}\right\}\right],\tag{22}$$

$$L=\bar{G}^{-1}\left[\exp\left\{\ln\bar{G}\left(x_m\right)-\left(a+W\right)\left(\frac{\alpha}{2}\right)^{1/(m+1)}\right\}\right],$$

其中，$\bar{G}^{-1}(\cdot)$ 表示 $\bar{G}(\cdot)$ 的反函数。当超参数 a 未知时，将前面给出的 a 的估计值 $\hat{a}$ 带入上式，即得逐步增加 II 型截尾样本 X_{m+1} 的经验 Bayes 预测区间。

4. 数值算例

本节以 Lomax 寿命分布作为比例危险率模型的特殊例子，给出该模型在平衡损失下逐步增加 II 型截尾样本的数值模拟，并研究估计结果的精确性。对于 Lomax 分布，其分布函数如下：

$$F\left(x\right)=1-\left(1+\frac{x}{\beta}\right)^{-\theta},x>0,$$

其中，$\bar{G}\left(x\right)=\left(1+\frac{x}{\beta}\right)^{-1},g\left(x\right)=\frac{1}{\beta}\left(1+\frac{x}{\beta}\right)^{-2}$。

利用文献[7]给出的算法，逐步增加 II 型截尾数据产生步骤如下：

(S1.) 产生 m 个均匀 $U(0,1)$ 分布的独立同分布随机变量 $W_1,W_2,\cdots,W_m$；

(S2.) 给定逐步增加 II 型截尾策略 $r_1,r_2,\cdots,r_m$，令 $V_i=W_i^{1/(i+r_m+r_{m-1}+\cdots+r_{m-i+1})},i=1,2,\cdots,m$；

(S3.) 再令 $U_i=1-\left(V_mV_{m-1}\cdots V_{m-i+1}\right),i=1,2,\cdots,m$，则 $U_1,U_2\cdots,U_m$ 是来自于均匀分布 $U(0,1)$ 的逐步增加 II 型截尾数据；

(S4.) 最后利用变换 $x_i=F^{-1}\left(U_i\right)$，得到来自任意分布 $F(\cdot)$ 的逐步增加 II 型截尾寿命试验数据，这里 $F^{-1}(\cdot)$ 表示分布 $F(\cdot)$ 的反函数。

下面利用 Monte-Carlo 数值模拟方法给出参数 θ、可靠性指标 $R(t),H(t)$ 的 Bayes 估计，具体步骤如下：

1) 对于给定超参数 a，利用参数 θ 的先验分布(7)产生 θ 的一组值，并选取其中一个当作参数 θ 的真值，记为 $\hat{\theta}$，将 $\hat{\theta}$ 带入(5)，得到 $R(t),H(t)$ 的真值；

2) 对给定的 $\hat{\theta},m,n,\left(r_1,r_2,\cdots,r_m\right)$，利用上面的方法，产生逐步增加 II 型截尾试验数据 $\left(x_1,x_2,\cdots,x_m\right)$；

3) 利用(18)给出超参数 a 的数值解 $\hat{a}$；

4) 对于给定的 t，由(10)、(13)给出指标的频率估计；由(15)、(17)和(18)得到平衡损失函数下指标的经验 Bayes 估计；利用(19)~(21)获得参数和可靠性指标的 HPD 区间估计；最后，利用(22)获得未来样本的预测区间。

当给定 $\beta=1,a=1.5,\theta=0.75,n=20,m=5$ 和 $r=(3,3,3,3,3)$ 时，产生如下的逐步增加 II 型截尾样本：0.0225、0.1192、0.1274、0.1358、1.2379。表 1 给出了参数及可靠性指标的 Bayes 估计的结果及预测区间。

对于 Bayes 点估计，表 1 中第一行表示预估计取 MLE 时得到的估计值，第二行表示预估计取 UMVU 估计时得到的估计值。同时，为了比较 HPD 区间估计的优劣性，表 1 中还给出了通常的 Bayes 置信(BCI)区间估计。这里，表 1 第一行表示区间估计的估计范围，第二行表示相应的区间长度。

为了检验估计的精度，做 5000 次重复模拟试验数据，利用下面的式子给出两种损失函数下各指标贝叶斯估计量的估计风险：

$$ER\left(\hat{\phi}\right)=\frac{1}{5000}\sum\left(\hat{\phi}-\phi\right)^2,$$

Table 1. Bayes estimation of $(\omega=0.5,c=1,t=1.1,\alpha=0.1)$ under balanced loss

表 1. 平衡损失下 Bayes 估计 $(\omega=0.5,c=1,t=1.1,\alpha=0.1)$

指标	θ	$R(t)$	$H(t)$
参数真值	0.5	0.6901	0.2381
MLE	1.0524	0.4580	0.5012
UMVU	0.8420	0.5071	0.4009
平衡均方估计	1.0332 0.7560	0.4810 0.5153	0.4777 0.3876
平衡 Linex 估计	0.6756 0.7379	0.7823 0.8412	0.4842 0.3669
HPD 区间	(0.3354, 1.5702) 1.2348	(0.2883, 0.7319) 0.4436	(0.1643, 0.7351) 0.5707
BCI 区间	(0.4181, 1.6855) 1.2674	(0.2862, 0.7335) 0.4473	(0.1993, 0.8025) 0.6032
预测区间		$x_6\in(2.3159,3.1172)$	

其中，$\hat{\phi}$ 表示真值 ϕ 相应的 Bayes 估计。表 2 给出了给出了当样本量不同时所采取的截尾策略，表 3、表 4 给出了当预估计分别取 MLE 和 UMVU 估计时的估计风险。

Table 2. Progressively increasing type-II censoring samples

表 2. 逐步增加 II 型截尾样本

样本量	$(r_1,r_2,\cdots,r_m)$
n = 20, m = 5	$(3,3,3,3,3)$
n = 20, m = 10	$(1,1,1,1,1,1,1,1,1,1)$
n = 30, m = 20	$(1,0,1,0,1,0,1,0,1,0,1,0,1,0,1,0,1,0,1,0)$

Table 3. Bayesian estimation risk of $(\omega=0.5,c=1,t=1.1)$ under balanced loss for exponential distribution

表 3. 指数分布在平衡损失下 Bayes 估计风险 $(\omega=0.5,c=1,t=1.1)$

(n,m)	$ER(\hat{\theta}_{BS})$	$ER(\hat{\theta}_{BL})$	$ER(\hat{R}_{BS})$	$ER(\hat{R}_{BL})$	$ER(\hat{H}_{BS})$	$ER(\hat{H}_{BL})$
(20, 5)	0.3114	0.1057	0.0918	0.1121	0.0769	0.0686
(20, 10)	0.1879	0.0429	0.0361	0.0641	0.0548	0.0382
(30, 15)	0.0634	0.0101	0.0086	0.0213	0.0111	0.0107

注：这里 δ_0 取相关指标的 MLE。

Table 4. Bayesian estimation risk of $(\omega = 0.7, c = -1, t = 0.9)$ under balanced loss for exponential distribution

表 4. 指数分布在平衡损失下 Bayes 估计风险 $(\omega = 0.7, c = -1, t = 0.9)$

(n,m)	$ER(\hat{\theta}_{BS})$	$ER(\hat{\theta}_{BL})$	$ER(\hat{R}_{BS})$	$ER(\hat{R}_{BL})$	$ER(\hat{H}_{BS})$	$ER(\hat{H}_{BL})$
(20, 5)	0.1077	0.0693	0.0533	0.0426	0.0317	0.0269
(20, 10)	0.0836	0.0500	0.0279	0.0253	0.0104	0.0136
(30, 15)	0.0312	0.0214	0.0085	0.0123	0.0063	0.0078

注：这里 δ_0 取相关指标的 UMVU。

下面结合上述 Monte-Carlo 模拟算例，对逐步增加 II 型截尾样本在比例危险率模型下的 Bayes 估计结果进行简要分析：

1) 比例危险率模型是一类应用广泛的寿命模型，在实际工业生产、可靠性理论等领域有着广泛的应用，通过 $\bar{G}(x)$ 选取可模拟多种失效模型。

2) 平衡损失是一类更加广泛的损失，通过对 $\rho(\gamma(\theta),\delta)$ 的选取，可以模拟各种损失情形，由数值模拟的结果可以看出，该损失比一般损失更加灵活；同时，权系数 ω 和 δ_0 为 $\gamma(\theta)$ 的估计提供了更大的选择空间，兼顾了 Bayes 方法和经典方法的优点；

3) 从表 1 可以看出，在平衡损失函数下获得的 Bayes 估计优于相应的预估计，这体现了平衡损失函数的灵活性和 Bayes 估计的优越性。同时，由模拟结果亦可以看到，当预估计较优时，在平衡损失函数下所得到 Bayes 估计大部分情况也较优。对于 Bayes 区间估计，通过比较区间长度可知，对于同样的置信水平 α，最大后验密度区间估计优于通常的 Bayes 置信区间估计。由表 3、表 4 可知，不同样本量和 c,t,ω 值下，Bayes 估计风险均随着样本的增大而减小，体现了频率的稳定性。最后，本文的研究尚有不足之处，比如，在实际操作中由于试验成本、研究人员以及试验设备等方面的限制，有时无法对所有的试验样品进行及时检测等，这些问题都在一定程度上影响推断精度。因此，在实验设计领域，未来的研究仍有广阔的空间等待深入探索。

参考文献

[1] Lawless, J.F. (2002) Statistical Models and Methods for Lifetime Data. Wiley. https://doi.org/10.1002/9781118033005

[2] Algarni, A., Elgarhy, M., M Almarashi, A., Fayomi, A. and R El-Saeed, A. (2021) Classical and Bayesian Estimation of the Inverse Weibull Distribution: Using Progressive Type-I Censoring Scheme. *Advances in Civil Engineering*, **2021**, Article ID: 5701529. https://doi.org/10.1155/2021/5701529

[3] Aljohani, H.M. (2021) Statistical Inference of Chen Distribution Based on Two Progressive Type-II Censoring Schemes. *Computers, Materials & Continua*, **66**, 2797-2814. https://doi.org/10.32604/cmc.2021.013489

[4] Balakrishnan, N. and Aggarwala, R. (2000) Progressive Censoring: Theory, Methods, and Applications. Birkhauser.

[5] Zellner, A. (1994) Bayesian and Non-Bayesian Estimation Using Balanced Loss Functions. In: Gupta, S.S. and Berger, J.O., Eds., *Statistical Decision Theory and Related Topics V*, Springer, 377-390. https://doi.org/10.1007/978-1-4612-2618-5_28

[6] Mohammad, J.J., Marchand, E. and Parsian, A. (2006) Bayes Estimation under a General Class of Balanced Loss Functions. https://api.semanticscholar.org/CorpusID:195817313

[7] Balakrishnan, N. and Sandhu, R.A. (1995) A Simple Simulational Algorithm for Generating Progressive Type-II Censored Samples. *The American Statistician*, **49**, 229-230. https://doi.org/10.1080/00031305.1995.10476150