第43卷　第1期
2024年3月
《新疆师范大学学报》(自然科学版)
Journal of Xinjiang Normal University
(Natural Sciences Edition)
Vol.43,No.1
Mar. 2024

Mlinex损失函数下反向帕累托分布形状参数的Bayes估计

何贵阳,周菊玲*
(新疆师范大学　数学科学学院,新疆　乌鲁木齐　830017)

摘　要:文章研究了Mlinex损失函数下反向帕累托分布的参数估计问题。在已知反向帕累托分布位置参数的情况下,给出形状参数的五种估计方法:极大似然估计、最大后验估计、经典Bayes估计、多层Bayes估计、E-Bayes估计,并推导出相应估计方法下的具体表达式。利用MC方法在R软件下进行数值模拟,对比模拟数据确定了参数估计的最优环境,并验证了估计方法的合理性和估计结果的准确性与稳健性,得到了E-Bayes估计为最优估计方法的结论;最后利用最优估计方法对实例进行数据拟合,确定了新疆县市级城市的人均城市道路面积可以利用反向帕累托分布近似拟合,并结合最终数据给出了相应的数据分析。

关键词: Mlinex损失函数;反向帕累托分布;E-Bayes估计;数值模拟;数据拟合
中图分类号:O212　　**文献标识码:**A　　**文章编号:**1008-9659(2024)01-0001-12

帕累托分布是一类经典的,能充分反映幂律特征的分布函数,因此一直备受关注。其名称是由意大利经济学家维弗雷多·帕累托(1848—1923)定义的,这一分布在经济学以外领域被称为“布拉德福分布”。因帕累托分布中变量的独特取值要求,即定义变量取值在特定数值之上,使得帕累托分布在应用上受限。如黄娟等人讨论了Pareto分布参数的经验Bayes(EB)单边检验问题,构造了参数的经验Bayes检验函数,证明了其具有渐近最优性,并且获得了收敛速度[1]。李超建等人介绍了基于帕累托分布的禽畜种苗交易系统入侵容忍模型,每台服务器的结构不同,但禽畜种苗交易网站服务内容相同,具有响应结果一致性[2]。温利民等人建立贝叶斯模型,讨论帕累托索赔额分布中参数的估计问题,得到了风险参数的极大似然估计、贝叶斯估计和信度估计,并证明了这些估计的强相合性[3]。钱小仕等人提到地震震级超过某一阈值的超出量分布可以近似为广义帕累托分布,并介绍了基于广义帕累托分布给出的若干地震活动性参数的估计公式[4]。张悦基于多种复杂删失数据研究帕累托分布的统计特性,构建了逐步II型删失下的拟合优度检验,推导了广义逐步II型删失下帕累托分布的寿命绩效指数的统计推断,讨论了适应性逐步II型删失模式下帕累托分布的竞争风险模型分析,并将结论推广到更复杂的广义指数分布[5]。通过查阅近些年关于帕累托分布的相关研究文献,可以确定帕累托分布只局限于刻画幂律特征的分布函数中的上尾部分[1-7]。从此特性出发,考虑变量的取值,如果将帕累托分布变量的定义域取相反数,则可以求解其受限外的区域,即刻画幂律特征的分布函数中的下尾部分。王超探讨了反向帕累托分布的统计推断问题,通过研究2010年我国655个城市人口规模,证明了中小型城市人口规模可以使用反向帕累托分布进行拟合[8]。简单的取值变化弥补了帕累托分布刻画区域的不足,同时也完善了对满足幂律特征区域的一种刻画问题。针对帕累托分布变量取相反数的特点,结合帕累托分布特点,提出反向帕累托分布。反向帕累托分布的密度函数和分布函数分别为

$$f(x;a,\lambda)=\lambda a^{-\lambda}x^{\lambda-1},0<x\leqslant a$$

$$F(x;a,\lambda)=a^{-\lambda}x^{\lambda},0<x\leqslant a$$

[收稿日期]2023-05-29　　[修回日期]2023-08-14
[作者简介]何贵阳(1999-),男,新疆石河子人,硕士研究生,主要从事数理统计方面研究。
*[通讯作者]周菊玲(1968-),女,新疆乌鲁木齐人,副教授,主要从事数理统计方面研究。

其中,a和λ分别为位置参数和形状参数,且$a > 0,\lambda > 0$,符号表示为$RP(a,\lambda)$.

在处理参数估计问题上,常见方法有极大似然估计、矩估计、熵估计、Bayes估计等。根本上是频率学派与贝叶斯学派就估计方法进行激烈讨论,其中频率学派的极大似然估计与贝叶斯学派的最大后验估计最具代表性。但无论是哪一学派的哪一种估计方法,都离不开样本信息与损失函数的选取,其中常用的损失函数有熵损失函数、平方损失函数、加权平方损失函数、Linex损失函数、复合Linex损失函数、Mlinex损失函数等。值得一提的是Mlinex损失函数,Mlinex损失函数是一类非对称损失函数,是由Podder在2004年提出的一种修正的线性指数损失函数,其具体表达式为

$$L(\lambda,\theta) = \omega\left[\left(\frac{\theta}{\lambda}\right)^{c} - c\ln\left(\frac{\theta}{\lambda}\right) - 1\right],\omega > 0, c \neq 0 \tag{1}$$

其中,θ是未知参数λ判别空间的一个估计。Mlinex损失函数虽是对原有损失函数的一种修正,但一直未停止对其进行研究。例如王琳等人基于逐步增加Ⅱ型截尾样本,研究了Mlinex损失下BurrⅫ部件可靠性指标的经验Bayes估计[9]。丁新月等人在Mlinex损失函数下,讨论了逆伽马分布尺度参数的Bayes估计及其可容许性[10]。李新鹏等人利用信度理论的方法得到了Mlinex损失函数下Bühlmann-Straub模型具有特殊相依效应的信度保费,进而推导出Mlinex损失函数下Bühlmann模型具有此种相依效应的信度保费[11]。

事先说明Mlinex损失函数中常数c的取值问题。文章只研究$c > 0$的情况,$c < 0$的情形类似,不做充分讨论。文章第一节为预备知识;第二节对频率学派极大似然估计与贝叶斯学派最大后验估计进行讨论,并推导反向帕累托分布形状参数在最大后验估计方法下的具体表达式;第三节介绍了在Mlinex损失函数下反向帕累托分布形状参数的经典Bayes估计,并推导出具体表达式;第四节在Mlinex损失函数下,讨论反向帕累托分布形状参数的多层Bayes估计与E-Bayes估计;第五节通过数值模拟,验证所列举估计方法的准确性、稳健性、可靠性;第六节在参数最优环境下,利用最优估计方法,进行数据拟合,确定新疆县市级城市的人均城市道路面积可以利用反向帕累托分布近似拟合,并结合最终数据给出相应的数据分析。

1 预备知识

在处理待估参数是客观存在但未知的一类估计问题时,常用的估计方法是经典频率学派观点下的极大似然估计(MLE)。

引理1[9] 若$X_1,X_2,\cdots,X_n$是来自$RP(a,\lambda)$分布的简单随机样本,其中a与λ分别为位置参数与形状参数。令$X = (X_1,X_2,\cdots,X_n)$,并且$x_1,x_2,\cdots,x_n$是其相应随机样本下的观察值,则$RP(a,\lambda)$中位置参数a与形状参数λ的极大似然估计分别为

$$\hat{\lambda}_{MLE} = n\left(\sum_{i=1}^{n}\ln\frac{\hat{a}_{MLE}}{x_i}\right)^{-1}$$

$$\hat{a}_{MLE} = \max\left\{x_i\right\} = x_{(n)}$$

在实际应用时,发现个别待估参数与样本有关,针对这类情况,贝叶斯学派提出了最大后验估计(MAP)方法。其估计原理是考虑待估参数的先验信息与样本信息有关,需选择合适的估计量使得后验分布密度达到最大值[12],同时最大后验估计作为Bayes估计的一种近似解,也有一定的研究价值。

2 形状参数λ的最大后验估计

由于参数λ的最大后验估计应使后验分布达到最大[12]。即选定合适的估计量$\hat{\lambda}$使$p(X|\lambda)\pi(\lambda)$达到最大,其中$\pi(\lambda)$是参数λ的先验分布密度,$p(X|\lambda)$是样本$X_1, X_2,\cdots, X_n$对参数λ的条件密度。从处理参数估计问题的原理上可以看出,极大似然估计是最大后验估计在$\pi(\lambda) \propto 1$的先验分布。接下来利用这一特点结合引理1给出的极大似然估计方法,推导出形状参数λ的最大后验估计方法。

定理1 若$X_1,X_2,\cdots,X_n$是来自$RP(a,\lambda)$分布的简单随机样本,其中a与λ分别为位置参数与形状参数。令$X = (X_1,X_2,\cdots,X_n)$,并且$x_1,x_2,\cdots,x_n$是其相应随机样本下的观察值,选取$\Gamma(\beta,\gamma)$作为形状参数λ的先验分布$\pi(\lambda)$,则在位置参数a已知的情况下,形状参数λ的最大后验估计为

$$\hat{\lambda}_{MAP}=\frac{n+\beta-1}{\gamma-t}$$

其中，$t=\sum_{i=1}^{n}(\ln x_i-\ln a)$，$\Gamma(\beta,\gamma)$中参数$\beta,\gamma$为超参数，且$\beta>0,\gamma>0$.

证明 选取形状参数λ的先验分布为$\Gamma(\beta,\gamma)$，其中参数β,γ为超参数，且$\beta>0,\gamma>0$，同时令$p(X|\lambda)$是样本$X_1,X_2,\cdots,X_n$对参数λ的条件密度，则有

$$\pi(\lambda)=\frac{\gamma^{\beta}}{\Gamma(\beta)}\lambda^{\beta-1}e^{-\gamma\lambda} \tag{2}$$

$$p(X|\lambda)=\prod_{i=1}^{n}\left(\lambda a^{-\lambda}x^{\lambda-1}\right)=\lambda^{n}a^{-n\lambda}\prod_{i=1}^{n}\left(x_i^{\lambda-1}\right)=\frac{\lambda^{n}}{\prod_{i=1}^{n}x_i}e^{\lambda t},t=\sum_{i=1}^{n}(\ln x_i-\ln a) \tag{3}$$

因此有

$$p(X|\lambda)\pi(\lambda)=\frac{\lambda^{n}}{\prod_{i=1}^{n}x_i}e^{\lambda t}\frac{\gamma^{\beta}}{\Gamma(\beta)}\lambda^{\beta-1}e^{-\gamma\lambda}=\frac{\gamma^{\beta}}{\Gamma(\beta)\prod_{i=1}^{n}x_i}\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda} \tag{4}$$

考虑到形状参数λ的最大后验估计是寻找λ的估计量$\hat{\lambda}$，使形状参数λ的后验密度函数达到最大值的情况，即找到$\hat{\lambda}$使$p(X|\lambda)\pi(\lambda)$达到最大值。

令$g(\lambda)=\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda}$，由最大后验估计方法的原理可知，要对$p(X|\lambda)\pi(\lambda)$关于$\lambda$求解最大值，就是要对$g(\lambda)$关于$\lambda$求解最大值。但关于$g(\lambda)$直接求解最值问题处理较为复杂，考虑变式，因$g(\lambda)=\exp\{\ln(g(\lambda))\}=\exp\{\ln(\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda})\}$，则对$g(\lambda)$求解最值问题可转化对$\ln[g(\lambda)]$求解最值问题。

对$\ln[g(\lambda)]$关于形状参数λ取一阶微分，同时令微商为0，即

$$\left(\ln[g(\lambda)]\right)'=\frac{n+\beta-1}{\lambda}-(\gamma-t)=0$$

解得形状参数λ的最大后验估计

$$\hat{\lambda}_{MAP}=\frac{n+\beta-1}{\gamma-t}$$

3 形状参数λ的Bayes估计

上文介绍了反向帕累托形状参数λ的最大后验估计。由于最大后验估计是Bayes估计解的近似值，其估计结果相比于利用Bayes理论下的经典估计方法得到的结果，还存在一定偏差。具体体现在损失函数的影响，所以下面将继续讨论在考虑损失函数情况下，反向帕累托形状参数λ在Bayes理论下的经典估计方法并确定估计结果的具体表达式。

考虑位置参数a已知的情况下，形状参数λ在Mlinex损失函数下的经典Bayes估计问题。

定理2 设$X_1,X_2,\cdots,X_n$是来自$RP(a,\lambda)$分布的简单随机样本，其中a与λ分别为位置参数与形状参数。令$X=(X_1,X_2,\cdots,X_n)$，并且$x_1,x_2,\cdots,x_n$是相应随机样本下的观察值，在Mlinex损失函数(1)下，对于任意的先验分布$\pi(\lambda)$，在位置参数已知的情况下，形状参数λ的唯一Bayes估计为

$$\hat{\lambda}_B(X)=\left(E\left(\lambda^{-c}|X\right)\right)^{-\frac{1}{c}}$$

证明 设$\hat{\lambda}(X)$为形状参数λ的任意一个估计，用$\hat{\lambda}(X)$代替损失函数中的θ，则在Mlinex损失函数下，$\hat{\lambda}(X)$相应的风险函数为

$$R_{\hat{\lambda}(X)}(\lambda)=E\left\{\omega\left[\left(\frac{\hat{\lambda}(X)}{\lambda}\right)^{c}-c\ln\left(\frac{\hat{\lambda}(X)}{\lambda}\right)-1\right]\right\}$$

但在贝叶斯观点下，$R_{\hat{\lambda}(X)}(\lambda)$是形状参数$\lambda$的函数，而$\lambda$还是随机变量，它也有先验分布$\pi(\lambda)$. 于是$\hat{\lambda}(X)$的损失函数应由$\int R_{\hat{\lambda}(X)}(\lambda)\pi(\lambda)d\lambda$判定，同时考虑贝叶斯观点下的最优估计，所以引入形状参数λ对X的“后验风

险”,即

$$R(\hat{\lambda}(X)|X)=\int\omega\left[\left(\frac{\hat{\lambda}(X)}{\lambda}\right)^{c}-c\ln\left(\frac{\hat{\lambda}(X)}{\lambda}\right)-1\right]p(X|\lambda)\pi(\lambda)d\lambda$$

其中,$p(X|\lambda)\pi(\lambda)$表示参数λ与样本 $X=(X_1,X_2,\cdots,X_n)$的联合密度函数。

结合贝叶斯观点下$\hat{\lambda}(X)$的损失函数,则有∫

$$\begin{aligned}\int R_{\hat{\lambda}(X)}(\lambda)\pi(\lambda)d\lambda&=\int E\left\{\omega\left[\left(\frac{\hat{\lambda}(X)}{\lambda}\right)^{c}-c\ln\left(\frac{\hat{\lambda}(X)}{\lambda}\right)-1\right]\right\}\pi(\lambda)d\lambda\\&=\underbrace{\int\cdots\int}_{n+1\text{次}}\omega\left[\left(\frac{\hat{\lambda}(X)}{\lambda}\right)^{c}-c\ln\left(\frac{\hat{\lambda}(X)}{\lambda}\right)-1\right]p(X|\lambda)\pi(\lambda)dXd\lambda\\&=\underbrace{\int\cdots\int}_{n\text{次}}R(\hat{\lambda}(X)|X)dX\\&=E\left\{E\left\{\omega\left[\left(\frac{\hat{\lambda}(X)}{\lambda}\right)^{c}-c\ln\left(\frac{\hat{\lambda}(X)}{\lambda}\right)-1\right]|X\right\}\right\}\end{aligned}$$

所以$\hat{\lambda}(X)$相应的风险函数为

$$\int R_{\hat{\lambda}(X)}(\lambda)\pi(\lambda)\,d\lambda=E\left\{E\left\{\omega\left[\left(\frac{\hat{\lambda}(X)}{\lambda}\right)^{c}-c\ln\left(\frac{\hat{\lambda}(X)}{\lambda}\right)-1\right]|X\right\}\right\}$$

其中,$E\left\{\omega\left[\left(\frac{\hat{\lambda}(X)}{\lambda}\right)^{c}-c\ln\left(\frac{\hat{\lambda}(X)}{\lambda}\right)-1\right]|X\right\}$表示参数$\lambda$与样本 $X=(X_1,X_2,\cdots,X_n)$的联合分布下的数学期望。

由损失函数定义可知,在对特定分布的参数进行估计时,考虑到给定相应损失函数后,需要使风险函数尽可能的小,以保证参数估计时的准确性。为此需使风险函数中的$E\left\{\omega\left[\left(\frac{\hat{\lambda}(X)}{\lambda}\right)^{c}-c\ln\left(\frac{\hat{\lambda}(X)}{\lambda}\right)-1\right]|X\right\}$极小化即可。

因为

$$E\left\{\omega\left[\left(\frac{\hat{\lambda}(X)}{\lambda}\right)^{c}-c\ln\left(\frac{\hat{\lambda}(X)}{\lambda}\right)-1\right]|X\right\}=\omega\left[(\hat{\lambda}(X))^{c}E\left(\lambda^{-c}|X\right)-c\ln(\hat{\lambda}(X))+cE\left(\ln\lambda|X\right)-1\right]$$

设
$$f(\hat{\lambda}(X))=(\hat{\lambda}(X))^{c}E\left(\lambda^{-c}|X\right)-c\ln(\hat{\lambda}(X))+cE\left(\ln\lambda|X\right)-1$$

将$f(\hat{\lambda}(X))$关于$\hat{\lambda}(X)$求一阶微分并令其等于零,便可解得形状参数λ的Bayes估计为

$$\hat{\lambda}_{B}(X)=\left(E\left(\lambda^{-c}|X\right)\right)^{-\frac{1}{c}}$$

由于$f(\hat{\lambda}(X))$是凸函数,所以$\hat{\lambda}(X)$是$f(\hat{\lambda}(X))$的唯一最小值。同时若存在λ'使得$R_{\hat{\lambda}(X)}(\lambda)<\infty$,对于参数$\lambda$的Bayes估计$\hat{\lambda}(X)$是唯一存在的且是可容许的,所以可以确定形状参数$\lambda$的唯一Bayes估计一般形式为

$$\hat{\lambda}_{B}(X)=\left(E\left(\lambda^{-c}|X\right)\right)^{-\frac{1}{c}}$$

推论1 同定理2条件。选取$\Gamma(\beta,\gamma)$作为$RP(a,\lambda)$分布中形状参数λ的先验分布$\pi(\lambda)$,其中参数β,γ为超参数,且$\beta>0,\gamma>0$,在Mlinex损失函数(1)下,且位置参数a已知的情况下,形状参数λ的Bayes估计的精确表达式为

$$\hat{\lambda}_{B^c}(X)=\left(E\left(\lambda^{-c}|X\right)\right)^{-\frac{1}{c}}=\frac{\left(\frac{\Gamma(n+\beta)}{\Gamma(n+\beta-c)}\right)^{\frac{1}{c}}}{(\gamma-t)}$$

证明 因为选取$\Gamma\left(\beta,\gamma\right)$作为形状参数$\lambda$的先验分布$\pi(\lambda)$，则由式(2)可得

$$\pi(\lambda)=\frac{\gamma^{\beta}}{\Gamma\left(\beta\right)}\lambda^{\beta-1}e^{-\gamma\lambda}$$

又因为$RP(a,\lambda)$分布的密度函数是$f(x;a,\lambda)=\lambda a^{-\lambda}x^{\lambda-1};0<x\leqslant a,\lambda>0$，所以样本的似然函数由式(3)确定为

$$L(X|\lambda)=\prod_{i=1}^{n}\left(\lambda a^{-\lambda}x^{\lambda-1}\right)=\lambda^{n}a^{-n\lambda}\prod_{i=1}^{n}\left(x_i^{\lambda-1}\right)=\frac{\lambda^{n}}{\prod_{i=1}^{n}x_i}e^{\lambda t},t=\sum_{i=1}^{n}(\ln x_i-\ln a)$$

因此，形状参数λ的后验分布密度为

$$\begin{aligned}\pi(\lambda|X)&=\frac{L(X|\lambda)\pi(\lambda)}{\int_0^{\infty}L(X|\lambda)\pi(\lambda)d\lambda}\\&=\frac{\frac{\lambda^{n}}{\prod_{i=1}^{n}x_i}e^{\lambda t}\frac{\gamma^{\beta}}{\Gamma\left(\beta\right)}\lambda^{\beta-1}e^{-\gamma\lambda}}{\int_0^{\infty}\frac{\lambda^{n}}{\prod_{i=1}^{n}x_i}e^{\lambda t}\frac{\gamma^{\beta}}{\Gamma\left(\beta\right)}\lambda^{\beta-1}e^{-\gamma\lambda}d\lambda}=\frac{\frac{\gamma^{\beta}}{\Gamma\left(\beta\right)\prod_{i=1}^{n}x_i}\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda}}{\frac{\gamma^{\beta}}{\Gamma\left(\beta\right)\prod_{i=1}^{n}x_i}\int_0^{\infty}\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda}d\lambda}\\&=\frac{\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda}}{\int_0^{\infty}\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda}d\lambda}=\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda}\frac{(\gamma-t)^{n+\beta}}{\Gamma(n+\beta)}\end{aligned}\tag{5}$$

由式(5)可以看出，形状参数λ的后验分布服从伽马分布$\Gamma(n+\beta,\gamma-t)$.

于是有

$$\begin{aligned}E\left(\lambda^{-c}|X\right)&=\int_0^{\infty}\lambda^{-c}\cdot\pi(\lambda|X)d\lambda=\int_0^{\infty}\lambda^{-c}\cdot\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda}\frac{(\gamma-t)^{n+\beta}}{\Gamma(n+\beta)}d\lambda\\&=\int_0^{\infty}\lambda^{n+\beta-c-1}e^{-(\gamma-t)\lambda}\frac{(\gamma-t)^{n+\beta}}{\Gamma(n+\beta)}d\lambda\\&=\frac{(\gamma-t)^{n+\beta}}{\Gamma(n+\beta)}\cdot\frac{\Gamma(n+\beta-c)}{(\gamma-t)^{n+\beta-c}}\\&=(\gamma-t)^{c}\frac{\Gamma(n+\beta-c)}{\Gamma(n+\beta)}\end{aligned}$$

因此，由定理2可知，Mlinex损失函数下形状参数λ的Bayes估计的精确表达式为

$$\hat{\lambda}_{B^c}(X)=\left((\gamma-t)^{c}\frac{\Gamma(n+\beta-c)}{\Gamma(n+\beta)}\right)^{-\frac{1}{c}}=\frac{\left(\frac{\Gamma(n+\beta)}{\Gamma(n+\beta-c)}\right)^{\frac{1}{c}}}{(\gamma-t)}$$

4 形状参数λ的E-Bayes估计(EB)与多层Bayes估计(HB)

在Bayes理论不断进步的同时，对特定分布参数的估计方法也一直不断地发展与完善。这一系列的发展也使得参数估计不断逼近于真值，使其误差不断地缩小，这样的结果正是对特定分布参数进行估计的最终理想。所以接下来文章进一步研究形状参数λ在Mlinex损失函数下，先验分布选定为$\Gamma\left(\beta,\gamma\right)$的E-Bayes估计与多层Bayes估计。根据相应文献，为了使估计的效果较好，$\Gamma\left(\beta,\gamma\right)$中参数$\beta$和$\gamma$的取值应使先验分布密度函数为形状参数$\lambda$的减函数[13]。再考虑估计的稳健性，最终确定$0<\beta<\gamma<m$，其中$m$为常数[14]。

4.1 E-Bayes估计(EB)

定义1 对于$(a,b)\in D$，若$\hat{\lambda}_B(a,b)$是连续的，则称

$$\hat{\lambda}_{EB}=\iint_D \hat{\lambda}_B(a,b)f(a,b)dadb$$

是参数λ的E-Bayes估计,其中$\iint_D \hat{\lambda}_B(a,b)f(a,b)dadb$是存在的,$D$是超参数$a$和$b$的取值集合,$f(a,b)$是$a$和$b$在集合$D$上的密度函数,$\hat{\lambda}_B(a,b)$为$\lambda$的Bayes估计。

从定义可以看出,参数λ的E-Bayes估计

$$\hat{\lambda}_{EB}=\iint_D \hat{\lambda}_B(a,b)f(a,b)dadb=E\left[\hat{\lambda}_B(a,b)\right]$$

是参数λ的Bayes估计$\hat{\lambda}_B(a,b)$对超参数a和b的数学期望,即λ的E-Bayes估计是λ的Bayes估计对超参数的数学期望。

定理3 设$X_1,X_2,\cdots,X_n$是来自$RP(a,\lambda)$分布的简单随机样本,其中a与λ分别为位置参数与形状参数。选取$\Gamma(\beta,\gamma)$作为形状参数λ的先验分布$\pi(\lambda)$,其中参数β,γ为超参数,且$\beta>0,\gamma>0$.令$X=(X_1,X_2,\cdots,X_n)$,并且$x_1,x_2,\cdots,x_n$是相应随机样本下的观察值,在位置参数已知时,$RP(a,\lambda)$分布中的形状参数λ,在Mlinex损失函数下的E-Bayes估计的精确表达式为

$$\hat{\lambda}_{EB}(X)=\frac{1}{m}\cdot\ln\frac{m-t}{-t}\cdot\int_0^1\left(\frac{\Gamma(n+\beta)}{\Gamma(n+\beta-c)}\right)^{\frac{1}{c}}d\beta$$

其中,$t=\sum_{i=1}^{n}(\ln x_i-\ln a)$,$\Gamma(\beta,\gamma)$中参数$\beta$和$\gamma$的先验分布分别为$U(0,1)$和$U(0,m)$.

证明 首先由推论1可知,$RP(a,\lambda)$分布的形状参数λ,在Mlinex损失函数下的Bayes估计的精确表达式为

$$\hat{\lambda}_{B^*}(X)=\frac{\left(\frac{\Gamma(n+\beta)}{\Gamma(n+\beta-c)}\right)^{\frac{1}{c}}}{(\gamma-t)}$$

其中,$t=\sum_{i=1}^{n}(\ln x_i-\ln a)$.给定形状参数$\lambda$的先验分布为$\Gamma(\beta,\gamma)$且$\beta,\gamma$均为超参数,$\beta>0,\gamma>0$.假设$\beta,\gamma$独立,则有$\beta$和$\gamma$的先验分布分别为$U(0,1)$和$U(0,m)$上的均匀分布,所以得到先验分布密度函数$f(\beta,\gamma)=\frac{1}{m}$.最后由定义1,$RP(a,\lambda)$分布的形状参数$\lambda$,在Mlinex损失函数下的E-Bayes估计的精确表达式为

$$\begin{aligned}\hat{\lambda}_{EB}(X)=E(\theta|X)&=\int_0^m\int_0^1\frac{\left(\frac{\Gamma(n+\beta)}{\Gamma(n+\beta-c)}\right)^{\frac{1}{c}}}{(\gamma-t)}f(\beta,\gamma)d\beta d\gamma\\&=\int_0^m\int_0^1\frac{\left(\frac{\Gamma(n+\beta)}{\Gamma(n+\beta-c)}\right)^{\frac{1}{c}}}{(\gamma-t)}\frac{1}{m}d\beta d\gamma\\&=\frac{1}{m}\int_0^1\left(\frac{\Gamma(n+\beta)}{\Gamma(n+\beta-c)}\right)^{\frac{1}{c}}d\beta\int_0^m\frac{1}{\gamma-t}d\gamma\\&=\frac{1}{m}\cdot\ln\frac{m-t}{-t}\cdot\int_0^1\left(\frac{\Gamma(n+\beta)}{\Gamma(n+\beta-c)}\right)^{\frac{1}{c}}d\beta\end{aligned}$$

4.2 多层Bayes估计(HB)

定义2 若λ的先验分布为$\Gamma(\beta,\gamma)$分布,其密度函数$\pi(\lambda)=\frac{\gamma^\beta}{\Gamma(\beta)}\lambda^{\beta-1}e^{-\gamma\lambda}$,其中参数$\beta,\gamma$为超参数,且$\beta>0,\gamma>0$.假设$\beta,\gamma$独立,则有$\beta$和$\gamma$的先验分布分别为$U(0,1)$和$U(0,m)$上的均匀分布,所以得到先验分布密度函数$f(\beta,\gamma)=\frac{1}{m}$,同时在$\beta$和$\gamma$独立时,则$\lambda$的多层先验密度函数为

$$\pi_*(\lambda)=\int_0^m\int_0^1\pi(\lambda)f(\beta,\gamma)d\beta d\gamma=\int_0^m\int_0^1\frac{\gamma^\beta}{\Gamma(\beta)}\lambda^{\beta-1}e^{-\gamma\lambda}\frac{1}{m}d\beta d\gamma,\lambda>0$$

定理4 同定理3条件。在位置参数a已知时，若形状参数λ的多层先验密度函数$\pi_*(\lambda)$由定义2给出，则在位置参数已知时，在Mlinex下形状参数λ的多层Bayes估计为

$$\hat{\lambda}_{HB}=\frac{\int_0^m\int_0^1\frac{\gamma^\beta\Gamma((n+\beta+1))}{\Gamma(\beta)(\gamma-t)^{n+\beta+1}}d\beta d\gamma}{\int_0^m\int_0^1\frac{\gamma^\beta\Gamma(n+\beta)}{\Gamma(\beta)(\gamma-t)^{n+\beta}}d\beta d\gamma}$$

其中，$t=\sum_{i=1}^n(\ln x_i-\ln a)$，$\Gamma(\beta,\gamma)$中参数$\beta$和$\gamma$的先验分布分别为$U(0,1)$和$U(0,m)$.

证明 设$X_1,X_2,\cdots,X_n$为来自$RP(a,\lambda)$分布的简单随机样本，在位置参数a已知时，样本的似然函数由式(3)给定

$$L(X|\lambda)=\prod_{i=1}^n(\lambda a^{-\lambda}x^{\lambda-1})=\lambda^n a^{-n\lambda}\prod_{i=1}^n(x_i^{\lambda-1})=\frac{\lambda^n}{\prod_{i=1}^n x_i}e^{\lambda t},t=\sum_{i=1}^n(\ln x_i-\ln a)$$

若形状参数λ的多层先验密度函数由定义2给出，根据Bayes定理，形状参数λ的多层后验分布密度为

$$\begin{aligned}h(\lambda|X)&=\frac{L(X|\lambda)\pi_*(\lambda)}{\int_0^\infty L(X|\lambda)\pi_*(\lambda)d\lambda}\\&=\frac{\frac{\lambda^n}{\prod_{i=1}^n x_i}e^{\lambda t}\cdot\int_0^m\int_0^1\frac{\gamma^\beta}{\Gamma(\beta)}\lambda^{\beta-1}e^{-\gamma\lambda}\cdot\frac{1}{m}d\beta d\gamma}{\int_0^\infty\frac{\lambda^n}{\prod_{i=1}^n x_i}e^{\lambda t}\cdot\int_0^m\int_0^1\frac{\gamma^\beta}{\Gamma(\beta)}\lambda^{\beta-1}e^{-\gamma\lambda}\cdot\frac{1}{m}d\beta d\gamma\,d\lambda}\\&=\frac{\int_0^m\int_0^1\frac{\gamma^\beta}{\Gamma(\beta)}\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda}d\beta d\gamma}{\int_0^m\int_0^1\frac{\gamma^\beta\Gamma(n+\beta)}{\Gamma(\beta)(\gamma-t)^{n+\beta}}d\beta d\gamma}\end{aligned}$$

在Mlinex损失函数下，形状参数λ的多层Bayes估计为

$$\begin{aligned}\hat{\lambda}_{HB}&=\int_0^\infty\lambda\cdot h(\lambda|X)d\lambda\\&=\int_0^\infty\lambda\cdot\frac{\int_0^m\int_0^1\frac{\gamma^\beta}{\Gamma(\beta)}\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda}d\beta d\gamma}{\int_0^m\int_0^1\frac{\gamma^\beta\Gamma(n+\beta)}{\Gamma(\beta)(\gamma-t)^{n+\beta}}d\beta d\gamma}d\lambda=\frac{\int_0^m\int_0^1\frac{\gamma^\beta}{\Gamma(\beta)}\left(\int_0^\infty\lambda\cdot\lambda^{n+\beta-1}e^{-(\gamma-t)\lambda}d\lambda\right)d\beta d\gamma}{\int_0^m\int_0^1\frac{\gamma^\beta\Gamma(n+\beta)}{\Gamma(\beta)(\gamma-t)^{n+\beta}}d\beta d\gamma}\\&=\frac{\int_0^m\int_0^1\frac{\gamma^\beta}{\Gamma(\beta)}\left(\int_0^\infty\lambda^{n+\beta}e^{-(\gamma-t)\lambda}d\lambda\right)d\beta d\gamma}{\int_0^m\int_0^1\frac{\gamma^\beta\Gamma(n+\beta)}{\Gamma(\beta)(\gamma-t)^{n+\beta}}d\beta d\gamma}=\frac{\int_0^m\int_0^1\frac{\gamma^\beta\Gamma((n+\beta+1))}{\Gamma(\beta)(\gamma-t)^{n+\beta+1}}d\beta d\gamma}{\int_0^m\int_0^1\frac{\gamma^\beta\Gamma(n+\beta)}{\Gamma(\beta)(\gamma-t)^{n+\beta}}d\beta d\gamma}\end{aligned}$$

5 数值模拟

文章研究了形状参数λ的五种估计方法并给出了相应的具体表达式。为确保估计所得结果的准确性、稳健性，接下来利用R软件对给出的估计方法进行MC数值模拟，并在模拟中运用控制变量的原理，观察对比偏差量Abs的数值变化，逐步得到最优估计的参数环境。在最优估计参数环境下，通过对均方误差MSE的数值变化进行讨论，最终确定Bayes理论下的最优估计。

在$RP(a,\lambda)$分布中，给定参数真值，即位置参数$a=100$和形状参数$\lambda=3$时，对样本取值为

$n=20$、50、100、150,Mlinex损失函数相应参数$\omega=1$、形状参数λ的先验分布为$\Gamma(2,1)$均给定。采用MC方法进行数值模拟计算,每种情况均进行2000次模拟计算,其计算结果如表1、表2、表3所示。其中,表1为给定条件下,确定Mlinex损失函数中常数c的最优环境;表2为给定条件下,确定形状参数λ的先验分布下参数的最优环境;表3为在最优环境下形状参数λ的三种估计方法下的均方误差MSE.

表1 确定Mlinex损失函数中常数c的最优环境(给定条件)

检验样本数与估计结果的相关性									
n	50			100			150		
MLE	3.1350			3.0580			3.0446		
(Abs)	0.1350			0.0580			0.0446		
MAP	2.9365			2.9595			2.9812		
(Abs)	0.0635			0.0405			0.0188		
Mlinex损失函数$\omega=1$,$n=50$、100、150,$c=1$、3、5,形状参数λ的先验分布为$\Gamma(2,1)$									
n	50			100			150		
c	1	3	5	1	3	5	1	3	5
B	2.9403	2.8720	2.8280	2.9637	2.9297	2.9043	2.9828	2.9628	2.9489
(Abs)	0.0597	0.1280	0.1720	0.0363	0.0703	0.0957	0.0172	0.0372	0.0511

由表1数据可得结论如下:

(1)与近似Bayes估计方法和Bayes估计方法相比,基本估计方法中极大似然估计结果比真值大,即偏差量为正。相比于极大似然估计,最大后验估计结果更接近真值。这是最大后验估计在估计问题上考虑了待估参数先验分布与样本相关的体现。

(2)Mlinex损失函数中常数c的取值变化对$RP(a,\lambda)$分布中形状参数λ的Bayes估计有一定的影响。可以看出,当Mlinex损失函数常数$c=1$时,其估计结果最接近真值,所以可以判断在已给定条件下,Mlinex损失函数参数的最优环境为$\omega=1$,$c=1$.

(3)对比三种估计方法在样本容量逐步增大时,估计结果也逐步趋近真值,即三种估计方法均满足大样本性质。同时对比偏差量可知所列举估计方法得到的结果是准确的,估计结果均满足准确性。

(4)结合数据可以判断出,在无损失函数环境影响下,处理满足反向帕累托分布形状参数的大样本数据的估计问题上,选择最大后验估计是最优的。

表2 确定形状参数λ的先验分布中参数的最优环境(给定条件)

n	m	B(Abs)	HB(Abs)	EB(Abs)
20	0.2	2.8308	3.1998(0.1998)	3.0323(0.0323)
	0.5		3.1065(0.1065)	2.9573(0.0427)
	1		3.0534(0.0534)	2.8556(0.1444)
50	0.2	2.9403 (0.0597)	3.0830(0.0830)	2.9626(0.0374)
	0.5		3.0457(0.0457)	2.9519(0.0481)
	1		3.0182(0.0182)	2.9318(0.0682)
100	0.2	2.9637 (0.0363)	3.0199(0.0199)	2,9871(0.0129)
	0.5		3.0144(0.0144)	2.9760(0.0240)
	1		3.0102(0.0102)	2.9575(0.0425)
150	0.2	2.9828 (0.0172)	3.0292(0.0292)	2.9866(0.1134)
	0.5		3.0165(0.0165)	2.9878(0.0122)
	1		3.0010(0.0010)	2.9714(0.0286)

注:$\omega=1$,$c=1$,$n=20$、50、100、150,形状参数λ的先验分布为$\Gamma(2,1)$且$m=0.2$、0.5、1.

由表2数据可得结论如下：

（1）在Bayes理论下，相比于经典Bayes估计（B）和E-Bayes估计（EB），多层Bayes估计（HB）估计结果的偏差为正偏差量，其他两个为负偏差量。

（2）形状参数λ的先验分布中参数的数值选取对Bayes估计结果有一定的影响，可以看出当先验分布中参数的数值选取$m = 1$时，其估计结果最接近真值。所以可以判断在已给定条件下，形状参数λ的先验分布中参数的数值选取最优环境是$m = 1$.同时对比m取值量的变化，三种Bayes估计方法得到结果变化的幅度都较小，即可说明三种Bayes估计方法所得结果均满足稳健性，其中E-Bayes估计稳健性最强。

（3）同表1结论，以上三种Bayes方法也均满足大样本性，估计结果均满足准确性。

（4）结合数据可以判断出：在Mlinex损失函数环境影响下，如需准确地处理满足反向帕累托分布形状参数的大样本数据估计问题时，选择E-Bayes估计方法是最优的。如需处理满足特定的正偏差逼近问题时，可选择多层Bayes估计方法得到较准确的近似值。

表3　在最优环境下形状参数λ的三种估计方法下的均方误差MSE

n	m	B	HB	EB
20		0.3551	0.4769	0.3640
50	1	0.1698	0.1859	0.1680
100		0.0882	0.0898	0.0882
150		0.0597	0.0602	0.0548

注：$\omega = 1, c = 1, n = 20$、50、100、150，形状参数$\lambda$的先验分布为$\Gamma(2,1)$且$m = 1$.

由表3数据可得结论如下：

（1）在参数的最优环境下，Bayes理论中的三种估计方法所得结果的均方误差MSE数值变化较稳定，结果能体现三种Bayes估计方法的合理可靠性。同时随着样本数据的增加，均方误差MSE均控制在较小的有效值内，即三种Bayes估计方法所得结果是准确有效的。

（2）数据对比可得：三种Bayes方法的均方误差MSE偏差量，在大样本数据下均控制在0.001，即三种Bayes估计方法所得结果较相近。同大样本数据下，E-Bayes估计方法中均方误差MSE控制较好，其可靠性较其他两种较强。

6　结论

文章所给出的包括近似Bayes与Bayes理论下常用的三种估计方法，通过数值模拟并分析比较，得到以上五种估计方法都满足大样本性质且部分估计结果具有一定的可靠性、准确性、稳健性。结合样本条件与稳健性要求，五种估计方法中E-Bayes估计法，在处理Mlinex损失函数下反向帕累托分布形状参数的估计问题上较为快捷、准确、稳定，即可判断E-Bayes估计是最优估计方法。

7　实例应用

文章1—4节内容已解决了所提出的估计问题，并最终做出总结，给出了处理相应参数估计问题的最优方法。但理论研究不仅仅是对处理问题的方法进行总结，更是要解决实际问题。在对反向帕累托分布的研究中，鲜有见到相关实际问题的讨论。王超探讨了反向帕累托分布的统计推断问题，通过2010年我国655个城市人口规模的案例，证明了中小型城市人口规模可以使用反向帕累托分布进行拟合[8]。蓝海等人基于E-Bayes估计的定义，分别在加权平方损失函数和平方损失函数下讨论了反向帕累托分布在位置参数已知时，形状参数α的E-Bayes估计[15]。徐宝等人使用加权p, q对称损失函数研究了反向帕累托分布的形状参数在刻度参数给定条件下Bayes估计的形式与性质。得到了形状参数Bayes估计的一般形式以及在给定共轭先验下的精确形式，证明了所得Bayes估计具有可容许性以及最小最大性[16]。文章将从文献[16]提出的反向帕累托分布可以拟合中小型城市人口规模的研究出发，对新疆维吾尔自治区二十座县市级城市的人均城市道路面积进行拟合研究。

一座城市的发展,不仅仅依赖于经济水平的提升,经济的发展与城市道路面积的扩建也体现在城市常住人口数的变化,但要考虑到经济发展同时伴随着人口流动。面对近些年不断发展的新疆,常住人口数已不能再作为衡量某座城市的发展标准,所以文章引入人均城市道路面积作为城市发展的判断依据。文章利用反向帕累托分布对新疆维吾尔自治区内二十座县市级城市的人均城市道路面积进行拟合,并利用最优估计方法判断拟合结果是否准确。以下数据来自于新疆维吾尔自治区统计局《2021年新疆统计年鉴11-2各城市市区设施水平》。

表4 2021年新疆20座县市级城市人均城市道路面积数据

城市名称	人均城市道路面积(m^2)	城市名称	人均城市道路面积(m^2)	城市名称	人均城市道路面积(m^2)
阿克苏市	12.69	哈密市伊州区	24.20	奎屯市	40.53
和田市	14.34	塔城市	24.55	吐鲁番市高昌区	46.01
博乐市	14.36	阿勒泰市	25.04	库尔勒市	52.52
昌吉市	15.42	乌苏市	27.98	克拉玛依市	55.54
伊宁市	16.97	库车市	28.49	阿拉山口市	72.99
乌鲁木齐市	19.68	阿图什市	30.58	霍尔果斯市	73.61
阜康市	23.72	喀什市	30.82		

注:表中数据不含兵团。

通过分析,设表中数据为 $X=(X_1,X_2,\cdots,X_{20})$ 的样本,通过计算得到样本均值与样本方差:$E(X)=32.50$,$Var(X)=345.50$.观察样本数据分布情况,有较多的小样本数据。同时结合人口分布的特点,数据可视为满足幂律特征的下尾分布,即考虑利用反向帕累托分布拟合。通过反向帕累托分布期望与方差公式

$$E(X)=\frac{\lambda a}{\lambda+1}$$

$$Var(X)=\frac{\lambda a^2}{(\lambda+1)^2(\lambda+2)}$$

结合样本均值与样本方差信息,求解得到参数真值 $a\approx 64.561$,$\lambda\approx 1.014$.但由于限定 $0<x\leqslant a$,所以上述表格中存在异常数据,从而限制了参数 a 的确定。对比实际人口数据判断阿拉山口市与霍尔果斯市数据存在一定异常,主要体现为流动人口数较多,常住人口数较其他地区偏少,使得在同等城市道路面积下,人均城市道路面积值偏大。

利用Excel软件,在理想环境下,对已知的20组数据进行参数的E-Bayes估计。通过迭代拟合,修正参数 a 的值并对估计结果与真值进行数值比较,在误差可允许范围内找到最优估计下参数的近似估计值,最后验证拟合的准确性。

由表5拟合结果可以得出:

(1)2021年新疆城市市区设施水平中人均城市道路面积数据可以用反向帕累托分布近似拟合,拟合结果相对准确。

(2)在处理2021年新疆城市市区设施水平中人均城市道路面积数据时,发现在给定数据的情况下,得到的参数 a 的真值存在误差,在后期数据拟合中,也验证了数据中阿拉山口市与霍尔果斯市数据存在异常。在数据不变的条件下,通过Excel软件的迭代修正参数 a 数值,并对修正后数据进行估计。对比参数 λ 真值,可以判断当 $a=244$ 时,估计结果与真值相同,数据拟合最完美。同时确定当 $a\in[136,845]$ 时,数据均方误差 $MSE\leqslant 0.3640$,即在可偏差范围内。

(3)在对参数 a 进行修正过程中发现,表中给定的数据中存在异常,但异常不是错误。根据对资料的查询与研究,找到阿拉山口市与霍尔果斯市数据异常原因为:该地区人口数据变化幅度较大,即流动人口数较多,常住人口数量较少,人口流动性较强。同时也说明该地区城市公共资源开发力度较强,开发后使用程度较低等问题。

(4)对比全国人居城市道路面积数据可以判断,以上城市中较多数城市数据高于全国标准数据17.36 m^2.

即说明新疆县市级城市资源利用率较低，固定人口数较少，人口流动性较强。

表5 通过迭代修正参数 a 的值并对参数 λ 拟合，得到近似拟合值 y（理想环境）

a		73.61	136	200	244	245	800	845	850
城市名称	人均城市道路面积（m^2）								
阿克苏市	12.69	−0.763	−1.030	−1.198	−1.284	−1.286	−1.800	−1.823	−1.826
和田市	14.34	−0.710	−0.977	−1.144	−1.231	−1.233	−1.747	−1.770	−1.773
博乐市	14.36	−0.710	−0.976	−1.144	−1.230	−1.232	−1.746	−1.770	−1.772
昌吉市	15.42	−0.679	−0.945	−1.113	−1.199	−1.201	−1.715	−1.739	−1.741
伊宁市	16.97	−0.637	−0.904	−1.071	−1.158	−1.159	−1.673	−1.697	−1.670
乌鲁木齐市	19.68	−0.573	−0.840	−1.007	−1.093	−1.095	−1.609	−1.633	−1.635
阜康市	23.72	−0.492	−0.758	0.926	−1.012	−1.014	−1.528	−1.551	−1.554
哈密市伊州区	24.20	−0.483	−0.750	0.917	−1.004	−1.005	−1.519	−1.543	−1.546
塔城市	24.55	−0.477	−0.743	0.911	−0.997	0.999	−1.513	−1.537	−1.539
阿勒泰市	25.04	−0.468	−0.735	0.902	−0.989	0.991	−1.504	−1.528	−1.531
乌苏市	27.98	−0.420	−0.687	0.854	−0.941	0.942	−1.456	−1.480	−1.483
库车市	28.49	−0.412	−0.679	0.846	−0.933	0.934	−1.448	−1.472	−1.475
阿图什市	30.58	−0.381	−0.648	0.816	−0.902	0.904	−1.418	−1.441	−1.444
喀什市	30.82	−0.378	−0.645	0.812	−0.899	0.900	−1.414	−1.438	−1.441
奎屯市	40.53	−0.259	−0.526	0.693	−0.780	0.781	−1.295	−1.319	−1.322
吐鲁番市高昌区	46.01	−0.204	−0.471	0.638	−0.725	0.726	−1.240	−1.264	−1.267
库尔勒市	52.52	−0.147	−0.413	0.581	−0.667	0.669	−1.183	−1.207	−1.209
克拉玛依市	55.54	−0.122	−0.389	0.556	−0.643	0.645	−1.158	−1.182	−1.185
阿拉山口市	72.99	−0.004	−0.270	0.438	−0.524	0.526	−1.040	−1.064	−1.066
霍尔果斯市	73.61	0	−0.267	−0.434	−0.520	−0.522	−1.036	−1.060	−1.062
T=sum（log（x）/a）		−8.321	−13.653	−17.002	−18.730	−18.765	−29.043	−29.519	−29.570
y=log（（1−l）*（−l）A（−1））*19.5		2.213	1.378	1.114	1.014	1.012	0.661	0.650	0.649

8 总结

文章对Mlinex损失函数下反向帕累托分布形状参数估计进行充分讨论，对比了频率学派的极大似然估计与贝叶斯学派的最大后验估计两大经典估计方法，两者的估计结果在数值上较为相似。结合两者在处理问题上的出发点不同，所以应用也各不相同。同时文章也在Bayes理论下，对相应参数进行了近似Bayes估计与经典Bayes估计的对比，给定了形状参数在估计时的参数最优环境，并通过数值模拟得到在Bayes理论下，处理相应估计问题的最优估计，即E-Bayes估计是最优估计方法。最后利用最优估计方法，对《2021年新疆统计年鉴11-2各城市市区设施水平》中人均城市道路面积的数据进行参数拟合，确定了新疆县市级城市的人均城市道路面积可以利用反向帕累托分布拟合，并结合最终数据给出了相应的数据分析。

参考文献：

[1] 黄娟，刘华祥，张健.φ混合样本下帕累托分布参数的经验贝叶斯分析[J].数学研究，2009，42(03)：335-340.

[2] 李超建，朱晓姝，龚榆桐.基于帕累托分布假设的禽畜种苗交易系统入侵容忍模型分析[J].沈阳农业大学学报，2014，45(01)：122-125.

[3] 温利民，张美，程子红，等.帕累托索赔分布中风险参数的经验贝叶斯估计[J].应用概率统计，2015，31(03)：225-237.

[4] 钱小仕,王福昌,盛书中.基于广义帕累托分布的地震震级分布尾部特征分析[J].地震学报,2013,35(03):341-350.
[5] 张悦.基于复杂删失数据帕累托分布的统计推断[D].北京:北京交通大学,2021.
[6] 陶瑞芳.帕累托分布中参数的线性贝叶斯估计[D].南昌:江西师范大学,2020.
[7] 石林.排序集抽样下帕累托分布形状参数的估计[D].武汉:华中师范大学,2014.
[8] 王超.反向帕累托分布参数估计及应用[J].统计与决策,2021,(14):41-44.
[9] 王琳,师义民,袁修国.MLINEX损失下BurrⅫ部件可靠性指标的经验贝叶斯估计[J].青岛科技大学学报(自然科学版),2011,32(02):204-207.
[10] 丁新月,徐美萍.Mlinex损失函数下逆伽马分布尺度参数的Bayes估计[J].广西师范大学学报(自然科学版),2014,(03):61-64.
[11] 李新鹏,吴黎军.MLINEX损失函数下具有风险相依效应的信度模型[J].江西师范大学学报(自然科学版),2020,44(01):17-19.
[12] 峁诗松,王静龙,濮晓龙.高等数理统计学[M].北京:高等教育出版社,1998.
[13] 韩明.多层先验分布的构造及其应用[J].运筹与管理,1997,6(03):31-40.
[14] BERGER J O. Statistical Decision Theory and Bayesian Analysis[M]. New York:Springer Verlag,1985.
[15] 蓝海,徐宝.两种平方损失下反向帕累托分布形状参数的E-Bayes估计[J].内江师范学院学报,2022,37(04):58-62.
[16] 徐宝,蓝海,赵仲达.一种对称损失函数下反向帕累托分布形状参数的估计[J].南开大学学报(自然科学版),2023,56(01):76-81.

Bayes Estimates of Shape Parameters of Reverse Pareto Distribution under Mlinex Loss Function

HE Gui-yang,ZHOU Ju-ling*
(*School of Mathematical Sciences,Xinjiang Normal University,Urumqi,Xinjiang,830017,China*)

Abstract: In this paper, the parameter estimation problem of inverted Pareto distribution under Mlinex loss function is studied. In the case of knowing the positional parameters of the inverted Pareto distribution, five estimation methods for the shape parameters are given: maximum likelihood estimation, maximum posterior estimation, classical Bayes estimation, multilayer Bayes estimation, E-Bayes estimation, and the specific expressions under the corresponding estimation methods are derived. The MC method is used to carry out numerical simulation under R software, the optimal environment for parameter estimation is determined by comparing the simulation data, and the rationality of the estimation method and the accuracy and robustness of the estimation results are verified, the conclusion that E-Bayes estimation is the optimal estimation method is obtained. Finally, the optimal estimation method is used to fit the data of the examples, and it is determined that the per capita urban road area of Xinjiang counties and cities can be approximated by using the reverse Pareto distribution, and the corresponding data analysis is given in combination with the final data.

Keywords: Mlinex loss function; Reverse Pareto distribution; E-Bayes estimation; Numerical simulation; Data fitting