文章编号：1674-8085(2017)03-0006-07

Pareto 分布在定时截尾样本下的估计问题

*刘荣玄[1]，邬四英[2]，王新长[1]

（1.井冈山大学数理学院，江西，吉安 343009 ；2 井冈山大学计划财务处，江西，吉安　343009）

摘　要：在定时截尾样本下研究两参数 Pareto 分布的参数估计，并计算它们的条件期望和方差；用数值模拟的办法验证所得估计的优劣，结果显示它们都是近似无偏估计(AUE)，且具有较好的稳定性；与相应的定数截尾样本下的估计相比较，它们的相对平均误差小。

关键词：两参数 Pareto 分布；定时截尾；近似无偏估计

中图分类号：O212.8　　**文献标识码**：A　　DOI:10.3969/j.issn.1674-8085.2017.03.002

PARAMETER ESTIMATION OF PARETO DISTRIBUTION UNDER TYPE Ⅰ CENSORING

* LIU Rong-xuan[1], WU Si-ying[2]，WANG Xin-chang[1]

(1. School of Mathematics and Physics, Jinggangshan University, Ji'an, Jiangxi 343009, China；

2. Department of Financial, Jinggangshan University,Ji'an,Jiangxi 343009,China)

Abstract: The parameter estimators of the shape and the scale parameter of Pareto distribution are studied under type Ⅰ censoring. The conditioning expectation and variance of these parameter estimators are calculated. The results of numerical simulation demonstrate that these estimators are asymptotic unbiased. The relative average errors of these estimators are lower than the counterpart under type Ⅱ censoring.

Key words: Pareto distribution; type Ⅰ censoring; asymptotic unbiased estimator

1　问题提出

1897 年意大利经济学家 Vilfredo Pareto (1848-1923)提出了Pareto分布[1]之后，Pareto分布得到非常广泛的应用，比如用于描述个人收入、城市人口容量、股票价格波动、江河流量等一些社会现象中隐含的客观规律。一个多世纪来，许多学者对两参数Pareto分布的参数估计问题进行了深入研究，取得一些好的成果，文献[2]在与信息论中的熵函数有关的q-对称熵损失函数下研究了Pareto分布形状参数的最小风险同变估计、Bayes估计以及可容许与不可容许估计，并给出了形状参数的一个最小最大估计；文献[3]基于逐步首失效样本，研究了Pareto分布可靠性指标的Bayes估计；文献[4]在逐步增加首失效截尾样本下，研究了参数估计的小样本性质和大样本性质。这些研究都是基于定数截尾样本进行的。在定时截尾样本下，有学者研究过指数分布的参数估计；文献[5]在定时截尾情形下对指数分布提出了修正的最大似然估计；文献[6]在定时截尾下研究了指数分布产品可靠性抽样检验方案的统计方法，并检验了统计量是平均寿命倒数的极大

收稿日期：2016-12-04；修改日期：2017-03-02
基金项目：江西省教育科学规划项目(16ZD026)
作者简介：*刘荣玄(1959－)，男，江西遂川人，副教授，主要从事概率论与数理统计教学与研究(E-mail：lrx1716@126.com);
邬四英(1963－)，女，江西丰城人，统计师，主要从事财务报表分析研究(E-mail：wusiying1963@163.com);
王新长(1971－)，男，江西于都人，副教授，博士，主要从事计量经济教学与研究(wangxinchang11@163.com).

似然估计；文献[7]在Bayes统计中基于定时截尾数据，研究了Pareto分布参数的Bayes估计。但是在经典统计中基于定时截尾样本研究Pareto分布参数的近似无偏估计(AUE)的文献没发现。基于此，本文着手于这项工作。在定时截尾试验中，由于在规定的时间$(0,\tau_0)$内产品失效的个数是一个随机变量，这给统计推断带来很大的麻烦，导致参数的无偏估计难以找到。有时一些学者不得不将定时截尾转化为定数截尾来处理，但这样一来又出现了信息丢失问题，具体来说，试验中最后一个失效产品的失效时间t_r与试验截尾时间τ_0之间没有产品失效的信息丢失了，影响估计的精度。基于这种情况，本文采用纠偏的方法，构造用于统计推断的统计量，给出两参数Pareto分布的AUE，并计算了估计量的条件期望和方差。

假设 Pareto 分布的概率密度函数为

$$f(x)=\begin{cases}\dfrac{1}{\alpha}\theta^{\frac{1}{\alpha}}x^{-\left(\frac{1}{\alpha}+1\right)}, & x\ge\theta\\ 0, & x<\theta\end{cases} \tag{1}$$

其中$\alpha>0$为形状参数，$\theta>0$为尺度参数，记为$X\sim Pareto(\theta,\alpha)$。

下文§2 将研究 Pareto 分布中形状参数α的 AUE，并计算它们的条件期望和方差，在§3 将研究 Pareto 分布中尺度参数θ的 AUE，并与最大似然估计和相应的定数截尾下的估计进行比较。

2 形状参数估计

为得到 Pareto 分布形状参数α的较为理想的 AUE，减少计算量，简化估计式，通过随机变量变换将参数α转到另一个分布函数中，进而研究它的估计问题。

假设随机变量$X\sim Pareto(\theta,\alpha)$，其概率密度函数为（1）式，令 $T=\ln X$, 则T 的概率密度函数为

$$f_T(t)=\begin{cases}\dfrac{1}{\alpha}e^{-\frac{1}{\alpha}(t-\mu)}, & t\ge\mu\\ 0, & t<\mu\end{cases} \tag{2}$$

其中$\mu=\ln\theta$。

从（2）式可知，T 服从参数为α，μ 的双参数指数分布，于是对 pareto 分布中参数α 的估计转移到对双参数指数分布中的α 估计，以简化计算。

在定时截尾试验中，假设截尾时间为τ_0，则样品在时间$(0,\tau_0)$内失效的概率为：

$$p=P\{X\le\tau_0\}=P\{e^T\le\tau_0\}=P\{T\le\tau\}=1-e^{-\frac{1}{\alpha}(\tau-\mu)},$$

其中$\tau=\ln\tau_0$， 令 $I=\begin{cases}1, & \mu\le T\le\tau,\\ 0, & T>\tau.\end{cases}$

于是有

$$E[(T-\tau)I]=E(TI)-\tau p=$$

$$\int_\mu^\tau t\frac{1}{\alpha}e^{-\frac{1}{\alpha}(t-\mu)}\mathrm{d}t-\tau p=-(\tau-\mu)+\alpha p=\alpha pg(p),$$

其中$g(p)=1+\dfrac{1}{p}\ln(1-p)$。

假设 $X_1,X_2,\cdots,X_n$ 是一列来自总体$X\sim Pareto(\theta,\alpha)$的容量为$n$的独立同分布(iid)样本，则由$X_1,X_2,\cdots,X_n$按$T_i=\ln X_i$生成的$T_1,T_2,\cdots,T_n$样本也是 iid 的，且服从双参数指数分布$Exp(\mu,\alpha)$。

根据矩估计思想，可用统计量$\frac{1}{n}\sum_{i=1}^{n}(T_i-\tau)I_i$作为$\alpha pg(p)$的无偏估计（U.E）。假设在$(0,\tau_0)$时间内有$r$个样品失效，显然它是一个随机变量，且$r\sim B(n,p)$。当失效时间$\tau_0$选择恰当时，在$(0,\tau_0)$内就会有一定比例的样品失效，根据频率收敛于概率的思想，可用失效频率$\frac{r}{n}$来近似替代失效概率p，于是得到参数α的第一次估计：

$$\hat{\alpha}_1=\frac{\sum_{i=1}^{n}(T_i-\tau)I_i}{rg(r/n)}$$

关于r的情况，当r=0 时，意味着在$(0,\tau_0)$时间内没有样品失效，$\frac{1}{n}\sum_{i=1}^{n}(T_i-\tau)I_i=0$，估计没有意义。当 r=n 时，意味着在$(0,\tau_0)$时间内所有投入试验的样品全部失效，所得到的样本是完全样本，有关完全样本的估计问题已有不少文献进行过讨

论。因此本文对在 $r\neq 0$，n 即 $0<r<n$ 的条件下，讨论 α 的估计问题。

在 $0<r<n$ 的条件下，$\hat{\alpha}_1$ 的数学期望为

$$E\left(\hat{\alpha}_1\middle|0<r<n\right)=E\left[\frac{\sum_{i=1}^{n}(T_i-\tau)I_i}{rg(r/n)}\middle|0<r<n\right]=$$

$$\frac{1}{1-p^n-q^n}\sum_{k=1}^{n-1}E_*\left[\frac{\sum_{i=1}^{n}(T_i-\tau)I_i}{kg(k/n)}\middle|r=k\right]p\{r=k\}=$$

$$\frac{1}{1-p^n-q^n}\sum_{k=1}^{n-1}\frac{C_n^k p^k q^{n-k}}{kg(k/n)}E_*\left[\sum_{i=1}^{n}(T_i-\tau)I_i\middle|r=k\right] \tag{3}$$

其中 $q=1-p$，E_* 表示对随机向量 $(T_1,T_2,\cdots,T_n)$ 求数学期望。因为样本 $T_1,T_2,\cdots,T_n$ 是 iid 的，所以有

$$E_*\left[\sum_{i=1}^{n}(T_i-\tau)I_i\middle|r=k\right]=kE_*\left[(T_i-\tau)I_i\middle|I_i=1\right]=k\alpha g(p) \tag{4}$$

将（4）式代入（3）式得

$$E\left(\hat{\alpha}_1\middle|0<r<n\right)=$$

$$\alpha g(p)\frac{1}{1-p^n-q^n}\sum_{k=1}^{n-1}\frac{C_n^k p^k q^{n-k}}{g(k/n)}=\alpha g(p)H_1(p) \tag{5}$$

其中 $H_1(p)=\dfrac{1}{1-p^n-q^n}\sum_{k=1}^{n-1}\dfrac{C_n^k p^k q^{n-k}}{g(k/n)}$。

根据文献[8]的作法，可将 $g(p)H_1(p)$ 看成修偏因子，用失效频率 $\frac{r}{n}$ 迭代失效概率 p，参数 α 的第一次估计 $\hat{\alpha}_1$ 迭代 $E(\hat{\alpha}_1|0<r<n)$，于是得到参数 α 的第二次估计为：

$$\hat{\alpha}_2=\frac{\hat{\alpha}_1}{g(p)H_1(p)}=\frac{\sum_{i=1}^{n}(T_i-\tau)I_i}{rg^2(r/n)H_1(r/n)}.$$

在 $0<r<n$ 条件下，$\hat{\alpha}_2$ 的数学期望为：

$$E\left(\hat{\alpha}_2\middle|0<r<n\right)=$$

$$E\left[\frac{\sum_{i=1}^{n}(T_i-\tau)I_i}{rg^2(r/n)H_1(r/n)}\middle|0<r<n\right]=$$

$$\frac{1}{1-p^n-q^n}\sum_{k=1}^{n-1}\frac{C_n^k p^k q^{n-k}}{kg^2(k/n)H_1(k/n)}$$

$$E_*\left[\sum_{i=1}^{n}(T_i-\tau)I_i\middle|r=k\right]=\alpha g(p)H_2(p)$$

其中 $H_2(p)=\dfrac{1}{1-p^n-q^n}\sum_{k=1}^{n-1}\dfrac{C_n^k p^k q^{n-k}}{g^2(k/n)H_1(k/n)}$.

依此类推可对参数 α 的估计进行 m 次迭代，得到参数 α 的第 m 次估计：

$$\hat{\alpha}_m=\frac{\sum_{i=1}^{n}(T_i-\tau)I_i}{rg^m(r/n)H_1(r/n)H_2(r/n)\cdots H_{m-1}(r/n)}.$$

其中 m 为有限正整数。

在 $0<r<n$ 条件下，它的数学期望为

$$E(\hat{\alpha}_m|0<r<n)=\alpha g(p)H_m(p),$$

其中 $H_m(p)=$

$$\frac{1}{1-p^n-q^n}\sum_{k=1}^{n-1}\frac{C_n^k p^k q^{n-k}}{g^m(k/n)H_1(k/n)H_2(k/n)\cdots H_{m-1}(k/n)}.$$

下面通过数值模拟说明这种迭代估计的效果。

当 $n=20$，$p=0.6$ 时，α 的估计 $\hat{\alpha}_m$ 随迭代次数 m 的增加其相对偏差的变化情况见图 1。

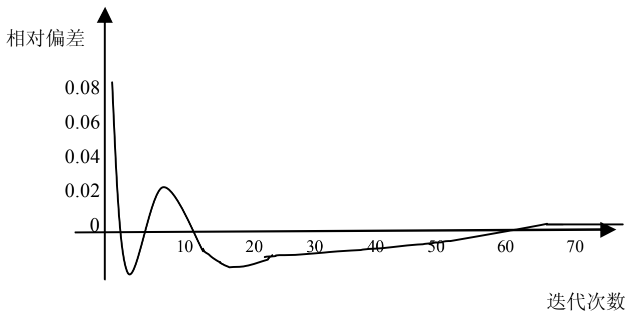

图 1 $n=20, p=0.6$ 时的迭代曲线

Fig.1 The iterative curve with parameter $n=20, p=0.6$

从图 1 可以看出，随着迭代次数 m 的增加，其相对偏差趋于稳定，且越来越接近于 0。从模拟实验中发现，对于 n,p 的其它组合也有同样的效果，只是所需迭代次数不同而已。

在不同失效概率 p 和不同样本容量 n 的情况下，迭代 20 次时，参数 α 估计 $\hat{\alpha}_{20}$ 的相对偏差均值的模拟结果见表 1。

表 1 $E\left(\frac{\hat{\alpha}_{20}-\alpha}{\alpha}\Big|0<r<n\right)$ 的值

Table 1 The value of $E\left(\frac{\hat{\alpha}_{20}-\alpha}{\alpha}\Big|0<r<n\right)$

p \ n	10	15	20	25	30	35	40	50
0.1	-0.3876	-0.0011	0.1690	0.2013	0.1532	0.0592	-0.0156	-0.0121
0.2	-0.0012	0.2044	0.0821	-0.0194	-0.0108	0.0180	0.0099	-0.0141
0.3	0.2184	0.0496	-0.0190	0.0224	0.0173	-0.0100	-0.0050	0.0092
0.4	0.2271	-0.0202	0.0408	-0.0009	-0.0165	0.0049	0.0040	-0.0055
0.5	0.1620	0.0463	0.0026	-0.0117	0.0100	0.0058	-0.0040	0.0020
0.6	-0.0423	0.0356	-0.0119	0.0108	0.0004	-0.0025	0.0011	-0.0009
0.7	-0.0956	-0.0105	0.0106	-0.0012	0.0051	0.0038	-0.0010	0.0004
0.8	0.0235	-0.0343	0.0098	-0.0059	0.0050	-0.0011	0.0006	0.0001
0.9	0.1967	0.0258	-0.0164	0.0009	-0.0023	0.0009	0.0007	0.0000

从表 1 可以看出，随着样本容量 n 和失效概率 p 的逐渐增大，平均相对偏差呈现逐步减小的趋势，且当 $n\ge 15,\ p\ge 0.3$ 时，估计 $\hat{\alpha}_{20}$ 的平均相对偏差不超过 5%，特别当样本容量 n 较大时，估计 $\hat{\alpha}_{20}$ 的平均相对偏差接近于零，即 $\hat{\alpha}_{20}$ 近似于参数 α 的无偏估计。

下面我们进一步讨论经过 m 次迭代所得估计 $\hat{\alpha}_m$ 的稳定性。

在 0＜r＜n 条件下， 估计 $\hat{\alpha}_m$ 的方差为：

$$Var\left(\hat{\alpha}_m\big|0<r<n\right)=E\left[\hat{\alpha}_m^2\big|0<r<n\right]-\left[E\left(\hat{\alpha}_m\big|0<r<n\right)\right]^2 \tag{6}$$

而

$$E\left[\left(\hat{\alpha}_m\right)^2\big|0<r<n\right]=E\left[\left(\frac{\sum_{i=1}^{n}\left(T_i-\tau\right)I_i}{rg^m\left(r/n\right)H_1\left(r/n\right)H_2\left(r/n\right)\cdots H_{m-1}\left(r/n\right)}\right)^2\Bigg|0<r<n\right]=$$

$$\frac{1}{1-p^n-q^n}\sum_{k=1}^{n-1}\frac{C_n^k p^k q^{n-k}}{\left[kg^m\left(k/n\right)H_1\left(k/n\right)H_2\left(k/n\right)\cdots H_{m-1}\left(k/n\right)\right]^2}E_*\left[\left(\sum_{i=1}^{n}\left(T_i-\tau\right)I_i\right)^2\Bigg|r=k\right] \tag{7}$$

因为

$$E_*\left[\left(\sum_{i=1}^{n}\left(T_i-\tau\right)I_i\right)^2\Bigg|r=k\right]=\sum_{i=1}^{n}E_*\left[\left(T_i-\tau\right)^2 I_i\big|r=k\right]+\sum_{i\ne l}E_*\left[\left(T_i-\tau\right)I_i\left(T_j-\tau\right)I_j\big|r=k\right] \tag{8}$$

由（8）式的第一个式子得：

$$\sum_{i=1}^{n}E_*\left[\left(T_i-\tau\right)^2 I_i\big|r=k\right]=\frac{k}{p}\int_{\mu}^{\tau}\left(t^2-2\tau t+\tau^2\right)\frac{1}{\alpha}e^{-\frac{1}{\alpha}(t-\mu)}\mathrm{d}t=\frac{k}{p}\left\{\left(\tau-\mu\right)^2-2\alpha\left(\tau-\mu\right)+2\alpha^2 p\right\}=k\alpha^2\left\{\frac{1}{p}\ln^2\left(1-p\right)+\frac{2}{p}\ln\left(1-p\right)+2\right\} \tag{9}$$

由（8）式第二个式子得：

$$\sum_{i\ne l}E_*\left[\left(T_i-\tau\right)I_i\left(T_j-\tau\right)I_j\big|r=k\right]=k\left(k-1\right)E_*\left[\left(T_i-\tau\right)I_i\left(T_j-\tau\right)I_j\big|I_i=1,I_j=1\right]=\alpha^2 k\left(k-1\right)g^2\left(p\right) \tag{10}$$

将（9），（10）两式代入（8）式得：

$$E_*\left[\left(\sum_{i=1}^{n}\left(T_i-\tau\right)I_i\right)^2\Bigg|r=k\right]=k\alpha^2\left\{\frac{1}{p}\ln^2\left(1-p\right)+\frac{2}{p}\ln\left(1-p\right)+\left(k-1\right)g^2\left(p\right)+2\right\} \tag{11}$$

将（11）式代入（7）式得：

$$E\left[\left(\hat{\alpha}_m\right)^2\big|0<r<n\right]=\frac{\alpha^2}{1-p^n-q^n}\cdot$$

$$\sum_{k=1}^{n-1}\frac{\left\{\frac{1}{p}\ln^2(1-p)+\frac{2}{p}\ln(1-p)+(k-1)g^2(p)+2\right\}C_n^k p^k q^{n-k}}{\left[g^m(k/n)H_1(k/n)H_2(k/n)\cdots H_{m-1}(k/n)\right]^2}$$

进而得到：

$$Var\left(\hat{\alpha}_m|0<r<n\right)=\frac{\alpha^2}{1-p^n-q^n}\cdot$$

$$\sum_{k=1}^{n-1}\frac{\left\{\frac{1}{p}\ln^2(1-p)+\frac{2}{p}\ln(1-p)+(k-1)g^2(p)+2\right\}C_n^k p^k q^{n-k}}{\left[g^m(k/n)H_1(k/n)H_2(k/n)\cdots H_{m-1}(k/n)\right]^2}-\alpha^2 g^2(p)H_m^2(p)$$

下面对迭代20次α的估计$\hat{\alpha}_m$在条件$0<r<n$下的相对方差进行数值计算，结果见表2。

表2 $Var\left(\frac{\hat{\alpha}_{20}}{\alpha}|0<r<n\right)$的值

Table 2 The value of $Var\left(\frac{\hat{\alpha}_{20}}{\alpha}|0<r<n\right)$

p \ n	10	15	20	25	30	35	40	50
0.1	0.2442	1.1869	2.4789	3.6263	4.3367	4.5579	3.9694	2.8898
0.2	1.1071	3.7471	4.6324	3.8790	2.7164	1.7789	0.7957	0.3971
0.3	2.5344	4.7091	3.1136	1.6037	0.8638	0.5029	0.1988	0.1340
0.4	4.0579	3.6111	1.4261	0.6443	0.3186	0.1306	0.0920	0.0918
0.5	5.0218	1.1253	0.6808	0.2721	0.1573	0.1305	0.0821	0.0658
0.6	4.9508	1.1253	0.3380	0.1687	0.1338	0.0970	0.0719	0.0568
0.7	3.8388	0.6931	0.2068	0.1527	0.1024	0.0860	0.0645	0.0514
0.8	2.1692	0.5619	0.2221	0.1282	0.1031	0.0850	0.0625	0.0499
0.9	0.7078	0.4446	0.2467	0.1931	0.1418	0.1006	0.0753	0.0373

从表2中可以看出，随着样本容量n和失效概率的逐渐增大，相对条件方差$Var\left(\frac{\hat{\alpha}_{20}}{\alpha}|0<r<n\right)$呈现逐步减少的趋势，特别当样本容量$n$较大时，估计$\hat{\alpha}_{20}$的相对条件方差越来越小，说明这种估计具有较好的稳定性。

3 尺度参数估计

下面将进一步研究 Pareto 分布尺度参数的AUE。本节所涉及的截尾时间τ_0，失效概率p和示性函数I均定义如前，即

$$p=\{X\le\tau_0\}=1-\theta^{\frac{1}{\alpha}}\tau_0^{-\frac{1}{\alpha}},\quad I=\begin{cases}1, & \theta\le X\le\tau_0,\\ 0 & X>\tau_0.\end{cases}$$

假设$X_1,X_2,\cdots,X_n$是一列来自总体分布（1）的容量为n的iid样本，由文献[9]知，总体的尺度参数θ的极大似然估计和定数截尾估计均为$X_{(1)}$，其中$X_{(1)}$为样本的最小次序统计量。

从（1）式可知，$X_{(1)}\ge\theta$，所以用$X_{(1)}$来估计θ估计值偏大，为达到纠偏的目的作如下处理。

因为$X_{(1)}$的概率密度函数为：

$$g\left(x_{(1)}\right)=C_n^1 f\left(x_{(1)}\right)\left[1-F\left(x_{(1)}\right)\right]^{n-1}=\frac{n}{\alpha}\theta^{\frac{n}{\alpha}}x_{(1)}^{-\left(\frac{n}{\alpha}+1\right)},x_{(1)}\ge\theta.$$

其数学期望为

$$E\left(X_{(1)}\right)=\int_\theta^\infty x\frac{n}{\alpha}\theta^{\frac{n}{\alpha}}x^{-\left(\frac{n}{\alpha}+1\right)}\mathrm{d}x=\frac{1}{1-\frac{\alpha}{n}}\theta,n>\alpha,$$

于是有

$$\theta=E\left(X_{(1)}\right)\left(1-\frac{\alpha}{n}\right),$$

用上一节所得到的α估计$\hat{\alpha}_m$的条件数学期望$E(\hat{\alpha}_m|0<r<n)$代替上式中的α，$X_{(1)}$代替$E(X_{(1)})$，于是得到参数θ的第m(m=1,2,⋯,)次迭代估计

$$\hat{\theta}_m=X_{(1)}\left(1-\frac{E\left(\hat{\alpha}_m|0<r<n\right)}{n}\right)\tag{12}$$

下面计算估计量$\hat{\theta}_m$在0<r<n下的数学期望，为此先证明下面几个结论。

引理1 假设$X_1,X_2,\cdots,X_n$是一列来自总体分布（1）的容量为n的iid样本，$X_{(1)}$为其最小次序统计量，r为产品在$(0,\tau_0)$时间内失效个数，则在0<r<n条件下，$X_{(1)}$的条件概率密度为

$$s(x)=\frac{\frac{n}{\alpha}\theta^{\frac{n}{\alpha}}\left[x^{-\left(\frac{n}{\alpha}+1\right)}-\left(x^{-\frac{1}{\alpha}}-\tau_0^{-\frac{1}{\alpha}}\right)^{n-1}x^{-\left(\frac{1}{\alpha}+1\right)}\right]}{1-p^n-q^n},\theta\le x\le\tau_0 .$$

证明：设随机变量$X_{(1)}$在条件$0<r<n$下的分布函数为$F(x)$，则对任意的$\theta\le x\le\tau_0$有

$$F(x)=P\left\{X_{(1)}\le x|0<r<n\right\}=\frac{P\left\{X_{(1)}\le x,0<r<n\right\}}{P\{0<r<n\}}=$$

$$\frac{P\left\{X_{(1)}\le x,0<r\le n\right\}-P\left\{X_{(1)}\le x,r=n\right\}}{P\{0<r<n\}} \quad (13)$$

$$P\left\{X_{(1)}\le x,0<r\le n\right\}=P\left\{X_{(1)}\le x\right\}=1-\theta^{\frac{n}{\alpha}}x^{-\frac{n}{\alpha}} \quad (14)$$

$$P\left\{X_{(1)}\le x,r=n\right\}=$$
$$P\left\{X_{(1)}\le x,X_1\le\tau_0,X_2\le\tau_0,\cdots,X_n\le\tau_0\right\}=$$
$$P\{X_1\le\tau_0,X_2\le\tau_0,\cdots,X_n\le\tau_0\}-$$
$$P\{x<X_1\le\tau_0,x<X_2\le\tau_0,\cdots,x<X_n\le\tau_0\}=$$
$$\left[1-\theta^{\frac{1}{\alpha}}\tau^{-\frac{1}{\alpha}}\right]^n-\left[\theta^{\frac{1}{\alpha}}\left(x^{-\frac{1}{\alpha}}-\tau_0^{-\frac{1}{\alpha}}\right)\right]^n \quad (15)$$

$$P\{0<r<n\}=1-p^n-q^n \quad (16)$$

将（14），（15），（16）式代入（13）式得

$$F(x)=\frac{1-\theta^{\frac{n}{\alpha}}x^{-\frac{n}{\alpha}}-\left[1-\theta^{\frac{1}{\alpha}}\tau_0^{-\frac{1}{\alpha}}\right]^n+\left[\theta^{\frac{1}{\alpha}}\left(x^{-\frac{1}{\alpha}}-\tau_0^{-\frac{1}{\alpha}}\right)\right]^n}{1-p^n-q^n}$$

进而可得其概率密度函数为

$$s(x)=F'(x)=\frac{\frac{n}{\alpha}\theta^{\frac{n}{\alpha}}\left[x^{-\left(\frac{n}{\alpha}+1\right)}-\left(x^{-\frac{1}{\alpha}}-\tau_0^{-\frac{1}{\alpha}}\right)^{n-1}x^{-\left(\frac{1}{\alpha}+1\right)}\right]}{1-p^n-q^n}$$

定理 1 在定时截尾寿命试验中假设截尾时间为τ_0，将n个服从分布（1）的样品$X_1,X_2,\cdots,X_n$全部投入试验，如果在时间$(0,\tau_0)$内有r个产品失效，则有

$$E\left[X_{(1)}|0<r<n\right]=\theta\frac{nq^n}{(1-p^n-q^n)}\cdot$$
$$\left[\frac{1}{n-\alpha}\left(q^{-n}-\frac{1}{q^\alpha}\right)+\sum_{m=0}^{n-1}\frac{C_{n-1}^m(-1)^{n-m}}{m+1-\alpha}\left(q^{-(m+1)}-\frac{1}{q^\alpha}\right)\right] \quad (17)$$

证明：因为

$$E\left[X_{(1)}|0<r<n\right]=$$
$$\int_\theta^{\tau_0}x\frac{\frac{n}{\alpha}\theta^{\frac{n}{\alpha}}\left[x^{-\left(\frac{n}{\alpha}+1\right)}-\left(x^{-\frac{1}{\alpha}}-\tau_0^{-\frac{1}{\alpha}}\right)^{n-1}x^{-\left(\frac{1}{\alpha}+1\right)}\right]}{1-p^n-q^n}\mathrm{d}x=$$
$$\frac{\frac{n}{\alpha}\theta^{\frac{n}{\alpha}}}{1-p^n-q^n}\left[\int_\theta^{\tau_0}x^{-\left(\frac{n}{\alpha}\right)}\mathrm{d}x-\int_\theta^{\tau_0}\left(x^{-\frac{1}{\alpha}}-\tau_0^{-\frac{1}{\alpha}}\right)^{n-1}x^{-\left(\frac{1}{\alpha}\right)}\mathrm{d}x\right]=$$
$$\frac{\frac{n}{\alpha}\theta^{\frac{n}{\alpha}}}{1-p^n-q^n}\cdot$$
$$\left[\frac{1}{\frac{n}{\alpha}-1}\left(\theta^{-\left(\frac{n}{\alpha}-1\right)}-\tau_0^{-\left(\frac{n}{\alpha}-1\right)}\right)-\int_\theta^{\tau_0}\sum_{m=0}^{n-1}C_{n-1}^m x^{-\frac{(m+1)}{\alpha}}\left(-\tau_0^{-\frac{1}{\alpha}}\right)^{n-1-m}\mathrm{d}x\right]=$$
$$\theta\frac{nq^n}{(1-p^n-q^n)}\cdot$$
$$\left[\frac{1}{n-\alpha}\left(q^{-n}-\frac{1}{q^\alpha}\right)+\sum_{m=0}^{n-1}\frac{C_{n-1}^m(-1)^{n-m}}{m+1-\alpha}\left(q^{-(m+1)}-\frac{1}{q^\alpha}\right)\right]$$

根据（12）式及定理 1 可知：

$$E\left(\hat{\theta}_m|0<r<n\right)=$$
$$E\left(X_{(1)}|0<r<n\right)\left(1-\frac{E\left(\hat{\alpha}_m|0<r<n\right)}{n}\right)=\theta\frac{nq^n}{(1-p^n-q^n)}$$
$$\left[\frac{1}{n-\alpha}\left(q^{-n}-\frac{1}{q^\alpha}\right)+\sum_{m=0}^{n-1}\frac{C_{n-1}^m(-1)^{n-m}}{m+1-\alpha}\left(q^{-(m+1)}-\frac{1}{q^\alpha}\right)\right].$$
$$\left(\frac{n-\alpha g(p)H_m(p)}{n}\right)$$

所以，

$$E\left(\frac{\hat{\theta}_m-\theta}{\theta}|0<r<n\right)=\frac{nq^n}{(1-p^n-q^n)}\cdot$$
$$\left[\frac{1}{n-\alpha}\left(q^{-n}-\frac{1}{q^\alpha}\right)+\sum_{m=0}^{n-1}\frac{C_{n-1}^m(-1)^{n-m}}{m+1-\alpha}\left(q^{-(m+1)}-\frac{1}{q^\alpha}\right)\right]\cdot$$
$$\left(\frac{n-\alpha g(p)H_m(p)}{n}\right)-1$$

再由（12）式得：

$$E\left(\frac{\hat{\theta}_m-\theta}{\theta}\middle|0<r<n\right)=$$

$$E\left[\left(\frac{X_{(1)}-\theta}{\theta}\middle|0<r<n\right)+1\right]\left(\frac{n-\alpha g(p)H_m(p)}{n}\right)-1$$

从而可得：

$$E\left(\frac{X_{(1)}-\theta}{\theta}\middle|0<r<n\right)=\frac{nq^n}{\left(1-p^n-q^n\right)}\cdot$$

$$\left[\frac{1}{n-\alpha}\left(q^{-n}-\frac{1}{q^{\alpha}}\right)+\sum_{m=0}^{n-1}\frac{C_{n-1}^{m}\left(-1\right)^{n-m}}{m+1-\alpha}\left(q^{-(m+1)}-\frac{1}{q^{\alpha}}\right)\right]-1$$

与极大似然估计和定数截尾估计比较，知道θ的极大似然估计和定数截尾估计均为$X_{(1)}$，且有$X_{(1)}\ge\theta$，所以$X_{(1)}$是θ的过高估计。为减少估计量，我们构造了$\hat{\theta}_m=X_{(1)}\left(1-\frac{E\left(\hat{\alpha}_m\middle|0<r<n\right)}{n}\right)$以达到纠偏之目的。

参考文献：

[1] Pareto V. Cours d'economie politique, Lausanne[J]. Tiattato di Socilolgia Generale, 1897.2:196

[2] 宋立新,王明秋,王晓光. q-对称熵损失函数下 Pareto 分布参数估计[J].大连理工大学学报,2011,51(4):616-620.

[3] 王亮,师义民. 两参数 Pareto 分布逐步首失效样本的 Bayes 估计[J].系统工程理论与践,2012,32(11): 2498-2503.

[4] 刘荣玄,吴高翔. 逐步增加首失效截尾样本下参数估计的优良性[J].应用概率统计,2015,31(2):135-145.

[5] 王炳兴,王玲玲. 定时截尾下指数分布的修正最大似然估计[J].高校应用数学学报,1995,10(3):295-302.

[6] 吴启光,吕建华. 定时截尾下指数分布产品可靠性抽样检验方案[J].系统科学与数学, 2003,23(2):145-154.

[7] 王晓红,宋立新. 定时截尾数据 Pareto 分布参数的 Bayes 估计[J].辽宁工程技术大学学报:自然科学版,2013,32(2):245-248.

[8] Mao Shisong, Han Qing. Statistical analysis of life and accelerated life test on weibull didtribution case under type 1 censoring [J]. Chinese Journal of Applied Probability and Statistics, 1991,7(1):61-72.

[9] Lawless J F. Statiscal Models and Methods for Lifetime data[M]. Ine：John Wiley & Sons,1982.