第 37 卷 第 1 期　　江西师范大学学报(自然科学版)　　Vol. 37 No. 1
2013 年 1 月　　Journal of Jiangxi Normal University(Natural Science)　　Jan. 2013

文章编号:1000-5862(2013)01-0016-04

缺失数据样本下艾拉姆咖分布的参数估计

龙　兵

(荆楚理工学院数理学院,湖北 荆门 448000)

摘要: 首先给出艾拉姆咖分布次序统计量的密度函数,得到次序统计量的相关特征数;其次在缺失数据样本下给出参数的无偏估计、近似极大似然估计及置信区间;最后通过实例计算出参数的几种估计,验证其优良性.

关键词: 艾拉姆咖分布;次序统计量;无偏估计;近似极大似然估计

中图分类号: O 212.1　　**文献标志码:** A

0　引言

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作用.国内对这类分布统计性质进行研究的文献较少.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.而实际上经常会遇到缺失数据样本,文献[4-8]在此类样本下进行了统计分析.在缺失数据样本下关于艾拉姆咖分布参数的估计问题鲜有研究,本文将由次序统计量的分布,给出在缺失数据样本下参数的点估计、置信区间,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

$$F(x)=1-(1+2x/\theta)\mathrm{e}^{-2x/\theta},\tag{1}$$

$$f(x)=\frac{4x}{\theta^2}\mathrm{e}^{-2x/\theta},\tag{2}$$

其中 $x>0$,参数 $\theta>0$.

由于 $E(X)=\theta$,在全样本场合下很容易得到参数 θ 的矩估计和极大似然估计,这里不作讨论.

引理 1[9]　设 $X_1,X_2,\cdots,X_n$ 独立同分布,有相同分布函数 $F(x)$ 和密度函数 $f(x)$,$X_{(1)},X_{(2)},\cdots,X_{(n)}$ 为其次序统计量,则第 i 个次序统计量 $X_{(i)}$ 的密度函数为

$$f_i(x)=\frac{n!}{(i-1)!(n-i)!}F^{i-1}(x)[1-F(x)]^{n-i}f(x),$$

特别地,$X_{(1)}$ 和 $X_{(n)}$ 的密度函数分别为

$$f_1(x)=n[1-F(x)]^{n-1}f(x),\ f_n(x)=nF^{n-1}(x)f(x).$$

1　次序统计量的性质

设 $X_{(1)},X_{(2)},\cdots,X_{(n)}$ 是来自艾拉姆咖分布(1)的 n 个次序统计量,则第 i 个次序统计量 $X_{(i)}$ 的密度函数为

$$f_i(x)=\frac{n!}{(i-1)!(n-i)!}[F(x)]^{i-1}[1-F(x)]^{n-i}f(x)=\frac{n!}{(i-1)!(n-i)!}\cdot\frac{4x}{\theta^2}\sum_{j=0}^{i-1}\binom{i-1}{j}(-1)^j\left(1+\frac{2x}{\theta}\right)^{n-i+j}\cdot\mathrm{e}^{-2(n-i+j+1)x/\theta}=\frac{4n!}{(i-1)!(n-i)!}\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\cdot\binom{n-i+j}{k}(-1)^j\frac{2^kx^{k+1}}{\theta^{k+2}}\mathrm{e}^{-2(n-i+j+1)x/\theta},x>0,\tag{3}$$

第 i 个次序统计量 $X_{(i)}$ 的 k(k 为正整数)阶原点矩为

$$E(X_{(i)}^k)=\frac{4n!}{(i-1)!(n-i)!}\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\binom{n-i+j}{k}\cdot(-1)^j\frac{2^k}{\theta^{k+2}}\int_0^{+\infty}x^{m+k+1}\mathrm{e}^{-2(n-i+j+1)x/\theta}\mathrm{d}x=$$

收稿日期: 2012-07-20

基金项目: 湖北省教育厅科研(B20102706)和荆楚理工学院院级科研(ZR201020)资助项目.

作者简介: 龙　兵(1973-),男,湖北荆门人,副教授,主要从事数理统计的研究.

$$\frac{n!\,\theta^m}{2^m(i-1)!(n-i)!}\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\binom{n-i+j}{k}\cdot(-1)^j\frac{(m+k+1)!}{(n-i+j+1)^{m+k+2}}.$$

当 $k=1$ 时,$X_{(i)}$ 的数学期望为

$$E(X_{(i)})=\frac{n!\,\theta}{2(i-1)!(n-i)!}\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\cdot\binom{n-i+j}{k}(-1)^j\frac{(k+2)!}{(n-i+j+1)^{k+3}}. \tag{4}$$

当 $k=2$ 时,$X_{(i)}$ 的 2 阶原点矩为

$$E(X_{(i)}^2)=\frac{n!\,\theta^2}{4(i-1)!(n-i)!}\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\binom{n-i+j}{k}(-1)^j\cdot\frac{(k+3)!}{(n-i+j+1)^{k+4}}.$$

因此,第 i 个次序统计量 $X_{(i)}$ 的方差为

$$D(X_{(i)})=\frac{n!\,\theta^2}{4(i-1)!(n-i)!}\cdot\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\binom{n-i+j}{k}(-1)^j\frac{(k+3)!}{(n-i+j+1)^{k+4}}-\left[\frac{n!\,\theta}{2(i-1)!(n-i)!}\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\binom{n-i+j}{k}\cdot(-1)^j\frac{(k+2)!}{(n-i+j+1)^{k+3}}\right]^2. \tag{5}$$

特别地,当 $i=1$ 时,最小次序统计量 $X_{(1)}$ 的密度函数为

$$f_1(x)=4n\sum_{k=0}^{n-1}\binom{n-1}{k}\frac{2^k}{\theta^{k+2}}x^{k+1}\mathrm{e}^{-2nx/\theta},x>0.$$

$X_{(1)}$ 的数学期望为

$$E(X_{(1)})=\frac{\theta}{2}\sum_{k=0}^{n-1}\binom{n-1}{k}\frac{(k+2)!}{n^{k+2}},$$

$X_{(1)}$ 的方差为

$$D(X_{(1)})=\frac{\theta^2}{4}\sum_{k=0}^{n-1}\binom{n-1}{k}\frac{(k+3)!}{n^{k+3}}-\frac{\theta^2}{4}\left[\sum_{k=0}^{n-1}\binom{n-1}{k}\frac{(k+2)!}{n^{k+2}}\right]^2.$$

当 $i=n$ 时,最大次序统计量 $X_{(n)}$ 的密度函数为

$$f_n(x)=4n\sum_{j=0}^{n-1}\sum_{k=0}^{j}\binom{n-1}{j}\binom{j}{k}(-1)^j\frac{2^kx^{k+1}}{\theta^{k+2}}\mathrm{e}^{-2(j+1)x/\theta},x>0.$$

$X_{(n)}$ 的数学期望为

$$E(X_{(n)})=\frac{n\theta}{2}\sum_{j=0}^{n-1}\sum_{k=0}^{j}\binom{n-1}{j}\binom{j}{k}(-1)^j\frac{(k+2)!}{(j+1)^{k+3}},$$

$X_{(n)}$ 的方差为

$$D(X_{(n)})=\frac{n\theta^2}{4}\sum_{j=0}^{n-1}\sum_{k=0}^{j}\binom{n-1}{j}\binom{j}{k}(-1)^j\frac{(k+3)!}{(j+1)^{k+4}}-\left[\frac{n\theta}{2}\sum_{j=0}^{n-1}\sum_{k=0}^{j}\binom{n-1}{j}\binom{j}{k}(-1)^j\frac{(k+2)!}{(j+1)^{k+3}}\right]^2.$$

2　θ 的估计

2.1　参数 θ 的近似极大似然估计

设 $X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n)}$ 为来自艾拉姆咖分布且容量为 n 的随机样本,由于某种原因导致部分数据缺失,只得到 k 个数据 $X_{(n_1)}\leqslant X_{(n_2)}\leqslant\cdots\leqslant X_{(n_k)}$,以上样本称为缺失数据样本(为方便起见,可将 $X_{(n_i)}$ 的下标数字省略括号).

令 $x=(x_{n_1},x_{n_2},\cdots,x_{n_k})$,则由文献[6]知,样本 x 的似然函数为

$$l(x\mid\theta)=C(F(x_{n_1}))^{n_1-1}(1-F(x_{n_k}))^{n-n_k}\prod_{i=2}^{k}(F(x_{n_i})-F(x_{n_{i-1}}))^{n_i-n_{i-1}-1}\prod_{i=1}^{k}f(x_{n_i}), \tag{6}$$

其中 C 为常数,$f(x)$ 与 $F(x)$ 为样本的密度函数与分布函数.

把(1)式和(2)式代入(6)式得到

$$l(x\mid\theta)=C\left[1-\left(1+\frac{2x_{n_1}}{\theta}\right)\mathrm{e}^{-2x_{n_1}/\theta}\right]^{n_1-1}\cdot\left[\left(1+\frac{2x_{n_k}}{\theta}\right)\mathrm{e}^{-2x_{n_k}/\theta}\right]^{n-n_k}\prod_{i=2}^{k}\left[\left(1+\frac{2x_{n_{i-1}}}{\theta}\right)\mathrm{e}^{-2x_{n_{i-1}}/\theta}-\left(1+\frac{2x_{n_i}}{\theta}\right)\mathrm{e}^{-2x_{n_i}/\theta}\right]^{n_i-n_{i-1}-1}\prod_{i=1}^{k}\frac{4x_{n_i}}{\theta^2}\mathrm{e}^{-2x_{n_i}/\theta},$$

对数似然函数为

$$\ln l(x\mid\theta)=\ln C+(n_1-1)\ln\left[1-\left(1+\frac{2x_{n_1}}{\theta}\right)\mathrm{e}^{-2x_{n_1}/\theta}\right]+(n-n_k)\ln\left(1+\frac{2x_{n_k}}{\theta}\right)-(n-n_k)\cdot\frac{2x_{n_k}}{\theta}+\sum_{i=2}^{k}(n_i-n_{i-1}-1)\ln\left[\left(1+\frac{2x_{n_{i-1}}}{\theta}\right)\mathrm{e}^{-2x_{n_{i-1}}/\theta}-\left(1+\frac{2x_{n_i}}{\theta}\right)\mathrm{e}^{-2x_{n_i}/\theta}\right]+\sum_{i=1}^{k}\ln 4x_{n_i}-2\sum_{i=1}^{k}\ln\theta-\sum_{i=1}^{k}\frac{2x_{n_i}}{\theta},$$

$$\frac{\partial\ln l(x\mid\theta)}{\partial\theta}=\frac{-(n_1-1)}{1-\left(1+\frac{2x_{n_1}}{\theta}\right)\mathrm{e}^{-2x_{n_1}/\theta}}\cdot\frac{4x^2}{\theta^3}\mathrm{e}^{-2x_{n_1}/\theta}-\frac{2(n-n_k)x_{n_k}}{\theta^2+2\theta x_{n_k}}+\frac{2(n-n_k)x_{n_k}}{\theta^2}+\sum_{i=2}^{k}\frac{n_i-n_{i-1}-1}{(\theta+2x_{n_{i-1}})\mathrm{e}^{-2x_{n_{i-1}}/\theta}-(\theta+2x_{n_i})\mathrm{e}^{-2x_{n_i}/\theta}}\cdot$$

$$\left(\frac{4x_{n_{i-1}}^2}{\theta^2}\mathrm{e}^{-2x_{n_{i-1}}/\theta}-\frac{4x_{n_{i-1}}^2}{\theta^2}\mathrm{e}^{-2x_{n_i}/\theta}\right)-\frac{2k}{\theta}+\frac{2}{\theta^2}\sum_{i=1}^{k}x_{n_i}.$$

令

$$\frac{\partial \ln l(x\mid\theta)}{\partial\theta}=0, \tag{7}$$

以 $\theta^{(j)}$ 表示 θ 的第 j 次迭代估计值,由(7)式得迭代公式

$$\theta^{(j+1)}=2k\left\{\frac{-(n_1-1)}{1-\left(1+\frac{2x_{n_1}}{\theta^{(j)}}\right)\mathrm{e}^{-2x_{n_1}/\theta^{(j)}}}\cdot\frac{4x_{n_1}^2}{(\theta^{(j)})^3}\mathrm{e}^{-2x_{n_1}/\theta^{(j)}}-\frac{2(n-n_k)x_{n_k}}{(\theta^{(j)})^2+2\theta^{(j)}x_{n_k}}+\frac{2(n-n_k)x_{n_k}}{(\theta^{(j)})^2}+\sum_{i=2}^{k}\frac{n_i-n_{i-1}-1}{(\theta^{(j)}+2x_{n_{i-1}})\mathrm{e}^{-2x_{n_{i-1}}/\theta^{(j)}}-(\theta^{(j)}+2x_{n_i})\mathrm{e}^{-2x_{n_i}/\theta^{(j)}}}\cdot\left(\frac{4x_{n_{i-1}}^2}{(\theta^{(j)})^2}\mathrm{e}^{-2x_{n_{i-1}}/\theta^{(j)}}-\frac{4x_{n_i}^2}{(\theta^{(j)})^2}\mathrm{e}^{-2x_{n_i}/\theta^{(j)}}\right)+\frac{2}{(\theta^{(j)})^2}\sum_{i=1}^{k}x_{n_i}\right\}^{-1},$$

将 θ 的最终迭代值记作 $\hat{\theta}_{AM}$,根据文献[10],则 $\hat{\theta}_{AM}$ 就是 θ 的近似极大似然估计(AMLE).

2.2 参数 θ 的无偏估计及区间估计

由(4)式,基于 $X_{(i)}$ 得到 θ 的1个无偏估计为

$$\hat{\theta}^{(i)}=2(i-1)!(n-i)!X_{(i)}\Big/\left[n!\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\binom{n-i+j}{k}(-1)^j\frac{(k+2)!}{(n-i+j+1)^{k+3}}\right], \tag{8}$$

则基于以上缺失数据样本可得到 θ 的1个无偏估计为

$$\hat{\theta}=\frac{1}{k}\sum_{i=1}^{k}\hat{\theta}^{(n_i)}=\frac{1}{k}\sum_{i=1}^{k}2(n_i-1)!(n-n_i)!X_{(n_i)}\Big/\left[n!\sum_{j=0}^{n_i-1}\sum_{k=0}^{n-n_i+j}\binom{n_i-1}{j}\binom{n-n_i+j}{k}(-1)^j\cdot\frac{(k+2)!}{(n-n_i+j+1)^{k+3}}\right].$$

引理2[11] 设 $\hat{\theta}_n=\hat{\theta}_n(x_1,x_2,\cdots,x_n)$ 是 θ 的1个估计量,若

$$\lim_{n\to\infty}E(\hat{\theta}_n)=\theta,\quad \lim_{n\to\infty}D(\hat{\theta}_n)=0,$$

则 $\hat{\theta}_n$ 是 θ 的相合估计.

引理3 若 $\hat{\theta}_{n1},\cdots,\hat{\theta}_{nk}$ 分别是 $\theta_1,\cdots,\theta_k$ 的相合估计,$\eta=g(\theta_1,\cdots,\theta_k)$ 是 $\theta_1,\cdots,\theta_k$ 的连续函数,则 $\hat{\eta}_n=g(\hat{\theta}_{n1},\cdots,\hat{\theta}_{nk})$ 是 η 的相合估计.

由(4)式和(8)式可得

$$E(\hat{\theta}^{(i)})=2(i-1)!(n-i)!E(X_{(i)})\Big/\left[n!\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\binom{n-i+j}{k}(-1)^j\cdot\frac{(k+2)!}{(n-i+j+1)^{k+3}}\right]=\theta,$$

由(5)式和(8)式可得

$$D(\hat{\theta}^{(i)})=[2(i-1)!(n-i)!]^2\cdot D(X_{(i)})\Big/\left\{n!\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\binom{n-i+j}{k}(-1)^j\cdot\frac{(k+2)!}{(n-i+j+1)^{k+3}}\right\}^2,$$

其中 $D(X_{(i)})$ 为(5)式所示.

要计算得到极限 $\lim\limits_{n\to\infty}D(\hat{\theta}^{(i)})=0$ 显然是很困难的,下面用数值方法通过计算机编程进行验证.

表1 $D(\hat{\theta}^{(i)})/\theta^2$ 的模拟值

	$n=60$	$n=100$	$n=1\,000$	$n=10\,000$	$n=100\,000$
$i=20$	0.020 03	0.018 08	0.014 25	0.006 23	0.000 72
$i=30$	0.015 05	0.013 02	0.009 76	0.003 06	0.000 54
$i=40$	0.013 07	0.010 49	0.007 46	0.002 13	0.000 41
$i=50$	0.013 42	0.009 03	0.006 08	0.001 12	0.000 23

从表1可以看出,对于某一固定的 i 值,随着 n 的增大,$D(\hat{\theta}^{(i)})/\theta^2$ 的值向0无限靠近.因此 $\hat{\theta}^{(i)}$ 是 θ 的相合估计,又由引理3可知 $\hat{\theta}$ 是 θ 的相合估计[12].

由(3)式,令 $X_{(i)}/\theta=U_i$,则 U_i 的密度函数为

$$f_{U_i}(u_i)=f_i(\theta u_i)\theta=\frac{4n!}{(i-1)!(n-i)!}\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\binom{n-i+j}{k}(-1)^j\cdot 2^k u_i^{k+1}\mathrm{e}^{-2(n-i+j+1)u_i},u_i>0. \tag{9}$$

由(9)式可以得到如下定理.

定理1 设 $X_1,X_2,\cdots,X_n$ 为来自艾拉姆咖分布且容量为 n 的随机样本,基于任一次序统计量 $X_{(i)}$,$i=1,2,\cdots,n$,可以得到 θ 的置信水平为95%的等尾置信区间为

$$\left[\frac{X_{(i)}}{d_i},\frac{X_{(i)}}{c_i}\right],$$

其中 c_i,d_i 满足

$$\int_0^{c_i}f_{U_i}(u_i)\,\mathrm{d}u_i=0.025,\int_{d_i}^{+\infty}f_{U_i}(u_i)\,\mathrm{d}u_i=0.025.$$

证 $0.95=1-\int_0^{c_i}f_{U_i}(u_i)\,\mathrm{d}u_i-\int_{d_i}^{+\infty}f_{U_i}(u_i)\,\mathrm{d}u_i=\int_{c_i}^{d_i}f_{U_i}(u_i)\,\mathrm{d}u_i=P(c_i\leqslant U_i\leqslant d_i)=P(c_i\leqslant X_{(i)}/\theta\leqslant d_i)=P(X_{(i)}/d_i\leqslant\theta\leqslant X_{(i)}/c_i)$,

以上置信区间的平均长度为

$$E\left(\frac{X_{(i)}}{c_i}-\frac{X_{(i)}}{d_i}\right)=\frac{d_i-c_i}{c_id_i}E(X_{(i)})=$$

$$\frac{n!(d_i-c_i)\theta}{2c_id_i(i-1)!(n-i)!}\sum_{j=0}^{i-1}\sum_{k=0}^{n-i+j}\binom{i-1}{j}\binom{n-i+j}{k}\cdot(-1)^j\frac{(k+2)!}{(n-i+j+1)^{k+3}}. \quad (10)$$

由于不能得到 c_i,d_i 的显示表达式,可用 Mathematica5.0 软件在计算机上计算得到近似值. 对于缺失数据样本 $X_{(n_1)}\leqslant X_{(n_2)}\leqslant\cdots\leqslant X_{(n_k)}$,可由(10)式得出 θ 的每一个置信区间的平均长度,取其中平均长度最短的置信区间作为 θ 的优化置信区间.

3 实例分析

通过 Monte-Carlo 模拟产生 1 组服从艾拉姆咖分布(参数真值 $\theta=3$) 的随机样本为

0.394 9,0.594 5,1.049 1,1.384 1,1.740 9,2.423 8,2.460 2,2.474 3,2.583 5,2.599 5,2.629 8,2.933 8,3.058 8,3.709 0,3.978 9,5.166 9,5.167 7,6.646 6,

根据以上完全样本可得到如下缺失数据样本:

缺失数据样本 1

0.594 5,1.049 1,1.740 9,2.460 2,2.474 3,2.583 5,2.629 8,2.933 8,3.709 0,5.166 9,6.646 6;

缺失数据样本 2

0.394 9,0.594 5,1.384 1,1.740 9,2.423 8,2.460 2,2.583 5,2.599 5,2.629 8,3.709 0,3.978 9,5.166 9,5.167 7;

缺失数据样本 3

0.394 9,0.594 5,1.049 1,1.740 9,2.474 3,2.583 5,2.599 5,3.058 8,3.709 0,3.978 9,5.167 7,6.646 6;

缺失数据样本 4

1.049 1,1.384 1,1.740 9,2.474 3,2.583 5,2.599 5,2.629 8,3.709 0,3.978 9,5.166 9;

缺失数据样本 5

0.394 9,0.594 5,1.740 9,2.423 8,2.460 2,2.474 3,2.583 5,2.599 5,2.629 8,2.933 8,3.058 8,3.978 9,5.166 9,6.646 6.

根据上面的结论,利用 Matlab 软件编程计算,可得到参数 θ 的几种估计.

表 2 参数 θ 的估计(置信水平 95%)

	完全样本	样本 1	样本 2	样本 3	样本 4	样本 5
$\hat{\theta}_{AM}$	2.833	2.895	2.863	2.764	2.946	2.854
$\hat{\theta}$	2.948	2.970	2.996	3.046	3.022	2.966
优化置信区间	[1.888,4.224]	[1.888,4.224]	[1.888,4.224]	[1.888,4.224]	[1.888,4.224]	[1.736,3.897]

通过上面的数据可以看出,在完全样本和不同的缺失数据样本下,所得到的参数 θ 的无偏估计 $\hat{\theta}$ 比近似极大似然估计 $\hat{\theta}_{AM}$ 更接近参数真值; 若设基于次序统计量 $X_{(i)}$ 所得到的 θ 置信区间的平均长度记为 l_i,则有

$l_{14}<l_{13}<l_{15}<l_{12}<l_{16}<l_{11}<l_{10}<l_{17}<l_9<l_8<l_7<l_{18}<l_6<l_5<l_4<l_3<l_2<l_1$,

因此,当 $n=18$ 时,$[X_{(14)}/d_{14},X_{(14)}/c_{14}]$ 是参数 θ 的最优化置信区间,从表 2 可知参数真值都落在 θ 的优化置信区间内部. 在完全样本场合下,文献[2]已经给出了求参数 θ 精确置信区间的方法,而对于缺失数据样本下参数的估计问题是不能解决的,本文的方法可以在任意程度的缺失数据样本下使用,且估计较为优良,它是对已有研究成果的一个补充.

4 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.

[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.

[3] 顾蓓青,王蓉华,徐晓岭. 艾拉姆咖分布的统计分析[C]//中国机械工程学会可靠性工程分会. 2011 年全国机械行业可靠性技术学会暨第 4 届可靠性工程分会第 3 次全体委员大会论文集. 北京:中国机械出版社,2011:65-67.

[4] 吕秋萍,邓文丽. 区间删失数据函数的均值估计[J]. 江西师范大学学报:自然科学版,2011,35(1):96-100.

[5] 游晓锋,丁树良,刘红云. 缺失数据的估计方法及应用[J]. 江西师范大学学报:自然科学版,2011,35(3):325-330.

[6] 龙兵,周良泽. 定数截尾数据缺失场合下冷贮备串联系统可靠性指标的经验 Bayes 估计[J]. 数学的实践与认识,2011,41(2):115-121.

[7] 任佳,高晓光,茹伟. 数据缺失的小样本条件下 BN 参数学习[J]. 系统工程理论与实践,2011,31(1):172-177.

[8] 官飞,邵敏,郭桐. 定时截尾样本数据有缺失情形下单参数对数正态分布的参数估计[J]. 阜阳师范学院学报:自然科学版,2009,26(4):29-32.

(下转第 27 页)

The Accuracy Measure for Rough Sets Based on Rank Statistics

WU Gen-xiu，LIU Pei-hong，LUO Bing-hui，XIE Jun
(College of Mathematics and Informatics，Jiangxi Normal University，Nanchang Jiangxi 330022，China)

Abstract: An accuracy measure for rough sets based on rank statistics has been proposed，which not only take into consideration the information of the size of the knowledge data value of particles but also the size of the universe. Meanwhile some properties are given. It is more reasonable and effective to measure，which are illustrated by examples.

Key words: rough sets; rank statistics; accuracy; improved knowledge capacity measure

(责任编辑：曾剑锋)

(上接第 19 页)

[9] 匡能晖. 关于两参数瑞利分布顺序统计量的分布性质[J]. 江西师范大学学报：自然科学版，2009，33(6)：648-651.

[10] 张莉. 分组数据在指数分布下的近似极大似然估计[J]. 统计与决策，2009(17)：153-155.

[11] 茆诗松，程依明，濮晓龙. 概率论与数理统计教程[M]. 2 版. 北京：高等教育出版社，2011.

[12] 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.

The Estimations of Parameter from Эрланга Distribution under Missing Data Samples

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

Abstract: Firstly，density function of order statistics is given about Эрланга distribution，calculate related characteristic numbers of the order statistics. Secondly，in the missing data samples unbiased estimation approximate maximum likelihood estimation and confidence interval of parameter are given. In the end，several estimations of the parameter are calculated through an instance，and its superiority is proved.

Key words: Эрланга distribution; order statistics; unbiased estimation; approximate maximum likelihood estimation

(责任编辑：曾剑锋)