文章编号：1004-4353(2018)03-0221-03

次序统计量的分布及其在均匀分布中的应用

时凌， 时义梅， 张琼
(广州工商学院 基础教学部，广东 广州 510850)

摘要： 讨论最大和最小次序统计量的分布以及它们在均匀分布中的应用，并讨论基于均匀分布 $U(\theta,\theta+1)$ 中的参数估计及优效性，得出 $\xi_{(1)}$，$\xi_{(n)}-1$ 是 θ 的均方相合估计和强相合估计的结论.

关键词： 次序统计量；均匀分布；参数估计
中图分类号： O212.1 **文献标识码：** A

Distribution of order statistics and its application in uniform distribution

SHI Ling， SHI Yimei， ZHANH Qiong
(*Basic Teaching Department*, *Guangzhou College of Technology and Business*, *Guangzhou* 510850, *China*)

Abstract: Maximum order statistics (or minimum order statistics) and its application in uniform distribution are discussed in this paper. Parametric estimation and its superiority for $U(\theta,\theta+1)$ uniform distribution are obtained.

Keywords: order statistics; uniform statistics; parametric estimation

次序统计量由于其独特的定义和性质在数理统计中受到人们的重视，并在参数估计的理论和计算中被广泛应用，特别是最大次序统计量和最小次序统计量. 文献[1-5]讨论了次序统计量的性质和分布、最大次序统计量和最小次序统计量的分布函数，以及最大次序统计量和最小次序统计量在参数估计中的应用. 本文在此基础上进一步讨论最大次序统计量和最小次序统计量在均匀分布 $U(\theta,\theta+1)$ 参数的估计及优效性问题，并得出了 $\xi_{(1)}$，$\xi_{(n)}-1$ 是 θ 的均方相合估计和强相合估计的结论.

1 次序统计量的定义

定义 1 设 $\xi_1,\xi_2,\cdots,\xi_n$ 为总体 ξ 的样本，由该样本建立 n 个函数：

$$\xi_{(k)}^*=\xi_{(k)}^*(\xi_1,\xi_2,\cdots,\xi_n),\ k=1,2,\cdots,n,$$

其中 $\xi_{(k)}^*$ 为该样本的统计量，它的观察值为 x_k^*，x_k^* 是样本 $\xi_1,\xi_2,\cdots,\xi_n$ 的观察值 $x_1,x_2,\cdots,x_n$ 由小到大排列(即 $x_1^*\leqslant x_2^*\leqslant\cdots\leqslant x_n^*$)后的第 k 位数值，则称 $\xi_{(1)}^*,\xi_{(2)}^*,\cdots,\xi_{(n)}^*$ 为次序统计量. 对于次序统计量 $\xi_{(1)},\xi_{(2)},\cdots,\xi_{(n)}$，显然有：

$$\xi_{(1)}=\min\{\xi_1,\xi_2,\cdots,\xi_n\},\ \xi_{(n)}=\max\{\xi_1,\xi_2,\cdots,\xi_n\},$$

分别称为最小次序统计量和最大次序统计量. 不难得到其密度函数分别为：

$$f_{\xi_{(1)}}(x)=n[1-F(x)]^{n-1}f(x),\ f_{\xi_{(n)}}(x)=n[F(x)]^{n-1}f(x),$$

收稿日期： 2018－06－27
作者简介： 时凌(1959—)，男，教授，研究方向为概率论与数理统计.

其中 $f(x)$ 和 $F(x)$ 分别称为总体 ξ 的密度函数和分布函数,并且$(\xi_{(1)},\xi_{(n)})$ 的联合密度函数为:

$$f(x,y)=\begin{cases}n(n-1)[F(y)-F(x)]^{n-2}f(x)f(y),\ (x,y)\in D;\\0,\ (x,y)\notin D.\end{cases}$$

其中 $D=\{(x,y)\mid y<x\}$.

2　均匀分布的次序统计量及其分布

假设总体 ξ 服从均匀分布 $U(\theta,\theta+1)$,则其密度函数和分布函数分别为:

$$f(x)=\begin{cases}1,\ \theta<x<\theta+1;\\0,\ 其它;\end{cases}$$

$$F(x)=\begin{cases}0,\ x<\theta;\\x-\theta,\ \theta\leqslant x<\theta+1;\\1,\ x>\theta.\end{cases}$$

其最小次序统计量和最大次序统计量的密度函数分别为:

$$f_{\xi_{(1)}}(x)=\begin{cases}n(1+\theta-x)^{n-1},\ \theta<x<\theta+1;\\0,\ 其它;\end{cases}$$

$$f_{\xi_{(n)}}(x)=\begin{cases}n(x-\theta)^{n-1},\ \theta<x<\theta+1;\\0,\ 其它.\end{cases}$$

$(\xi_{(1)},\xi_{(n)})$ 的联合密度函数为:

$$f(x,y)=\begin{cases}n(n-1)(y-x)^{n-2},\ (x,y)\in D;\\0,\ (x,y)\notin D.\end{cases}$$

其中 $D=\{(x,y)\mid y<x\}$.

3　均匀分布 $U(\theta,\theta+1)$ 的参数估计

3.1　矩估计

因为 $E(\xi)=\int_{\theta}^{\theta+1}x\,\mathrm{d}x=\theta+\frac{1}{2}$,令 $\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i$,从而 $\bar{\xi}=\hat{\theta}+\frac{1}{2}$,由此得 $\hat{\theta}=\bar{\xi}-\frac{1}{2}$ 是 θ 的矩估计,且还是 θ 的无偏估计.

3.2　极大似然估计

参数 θ 的似然函数为 $L(\theta)=\begin{cases}1,\ \theta<x_{(1)},x_{(n)}<\theta+1;\\0,\ 其它.\end{cases}$ 因为 $\theta<x_{(1)},x_{(n)}<\theta+1$,所以 $x_{(n)}-1<\theta<x_{(1)}$.由此知 $\hat{\theta}_1=\min\{\xi_1,\xi_2,\cdots,\xi_n\}=\xi_{(1)}$ 和 $\hat{\theta}_2=\max\{\xi_1,\xi_2,\cdots,\xi_n\}-1=\xi_{(n)}-1$ 均为 θ 的极大似然估计(MLE),且 $(1-\lambda)\xi_{(1)}+\lambda\xi_{(n)}$ 也是 θ 的极大似然估计(MLE),所以 θ 的极大似然估计不惟一,其中 $0<\lambda<1$.

3.3　最小方差无偏估计

$$D(\hat{\theta})=D\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{1}{2}\right)=D\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)=D(\bar{\xi})=\frac{1}{n}D(\xi)=\frac{1}{n}(E(\xi^2)-(E(\xi))^2)=$$
$$\frac{1}{n}\left[\int_{\theta}^{\theta+1}x^2\,\mathrm{d}x-(\theta+\frac{1}{2})^2\right]=\frac{1}{12n},$$

$$E(\xi_{(1)})=\int_{\theta}^{\theta+1}nx(1+\theta-x)^{n-1}\,\mathrm{d}x=\int_{0}^{1}n(1+\theta-y)y^{n-1}\,\mathrm{d}y=n\left(\frac{1+\theta}{n}-\frac{1}{n+1}\right)=\theta+\frac{1}{n+1},$$

$$D(\xi_{(1)})=E(\xi_{(1)}^2)-(E(\xi_{(1)}))^2=\int_{\theta}^{\theta+1}nx^2(1+\theta-x)^{n-1}\,\mathrm{d}x-\left(\theta+\frac{1}{n+1}\right)^2=$$

$$n\int_0^1 (1+\theta-y)^2 y^{n-1}\mathrm{d}y-\left(\theta+\frac{1}{n+1}\right)^2=\frac{n}{(n+1)^2(n+2)},$$

$$E(\xi_{(n)}-1)=E(\xi_{(n)})-1=\int_\theta^{\theta+1} nx(x-\theta)^{n-1}\mathrm{d}x-1=\int_0^1 n(y+\theta)y^{n-1}\mathrm{d}y-1=$$

$$\theta+\frac{n}{n+1}-1=\theta-\frac{1}{n+1},$$

$$E(\xi_{(n)})=\theta+1-\frac{1}{n+1}=\theta+\frac{n}{n+1},$$

$$D(\xi_{(n)}-1)=D(\xi_{(n)})=E(\xi_{(n)}^2)-(E(\xi_{(n)}))^2=\int_\theta^{\theta+1} nx^2(x-\theta)^{n-1}\mathrm{d}x-\left(\theta+\frac{n}{n+1}\right)^2=$$

$$\int_0^1 n(\theta+y)^2 y^{n-1}\mathrm{d}y-\left(\theta+\frac{n}{n+1}\right)^2=\frac{n}{(n+1)^2(n+2)}.$$

又因为当 $n>10$ 时,有$\frac{1}{12n}>\frac{n}{(n+1)^2(n+2)}$,所以当样本量充分大时 $\xi_{(1)},\xi_{(n)}-1$ 比 $\hat{\theta}$ 更优.

4 均匀分布 $U(\theta,\theta+1)$ 的参数估计量的相合性

定理 1 假设$(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}$为来自均匀分布$U(\theta,\theta+1)$的样本,则:1) $\xi_{(1)},\xi_{(n)}-1$ 是 θ 的均方相合估计;2) $\xi_{(1)},\xi_{(n)}-1$ 是 θ 的强相合估计.

证明 因为

$$E|\xi_{(1)}-\theta|^2=E(\xi_{(1)}-\theta)^2=E(\xi_{(1)}^2-2\theta\xi_{(1)}+\theta^2)=E(\xi_{(1)}^2)-2\theta E(\xi_{(1)})+\theta^2=$$

$$(1+\theta)^2-\frac{2n(1+\theta)}{n+1}+\frac{n}{n+2}-2\theta\left(\theta+\frac{1}{n+1}\right)+\theta^2=\frac{2}{(n+1)(n+2)}\to 0\ (n\to\infty),$$

根据均方相合估计定义知 $\xi_{(1)}$ 是 θ 的均方相合估计.同理可证 $\xi_{(n)}-1$ 也是 θ 的均方相合估计.

又因为 $P(|\xi_{(1)}-\theta|\geqslant\varepsilon)=\int_{|x-\theta|\geqslant\varepsilon}\mathrm{d}F(x)\leqslant\frac{D(\xi_{(1)})}{\varepsilon^2}=\frac{1}{\varepsilon^2}\frac{n}{(n+1)(n+2)}\to 0\ (n\to\infty)<+\infty$,由 Borel-Cantelli 引理得 $\xi_{(1)}$ 是 θ 的强相合估计.同理可证 $\xi_{(n)}-1$ 也是 θ 的强相合估计.证毕.

参考文献:

[1] 熊加兵,王志祥.均匀分布次序统计量的性质[J].高等数学研究,2010,13(1):17-19.

[2] 孙婷婷,洪义成,沈京虎,等.次序统计量概论密度函数的新的推导方法[J].延边大学学报(自然科学版),2016,42(3):192-195.

[3] 沈伶伶.顺序统计统计量的分布及在优铲性中的应用[J].湖北民族学院学报(自然科学版),2012,30(2):186-188.

[4] 何朝兵,田彦伟.顺序统计统计量的分布[J].成都大学学报(自然科学版),2008,27(2):116-119.

[5] 胡林娟.特色分布下的次序统计量的性质研究[D].杭州:浙江大学,2016.

[6] 梁之舜,邓集贤,杨维权,等.概率论及数理统计:上、下册[M].北京:高等教育出版社,2009.

[7] 盛骤,谢式千,潘承毅.概率论与数理统计[M].北京:高等教育出版社,2001.

[8] 李贤平.基础概率论[M].北京:高等教育出版社,2010.

[9] 范金城,吴可法.统计推断导引[M].北京:科学出版社,2001.