第12卷 第1期 郑州工学院学报 Vol.12 No.1
1991年 3 月 Journal of Zhengzhou Institute of Technology Mar. 1991

改进的对称分布期望估计量的性质*

张新育

（数理力学系）

摘 要：本文证明了改进后的对称分布期望估计量 W_p 在一定条件下具有无偏性和强相合性，并给出了方差的上界，其中：

$$W_{p}=\frac{1}{n-2[np]}\sum_{i=[np]+1}^{n-[np]}X_{(i)},\quad 0\leqslant p<\frac{1}{2}$$

关键词：数学期望，估计，无偏性

中国图书分类号：0211:0212

在实践中人们为避免细尾对称分布的异常值对均值的影响，常采用去掉一个最小值、一个最大值或同时剔除几次的办法．如体操比赛成绩的产生就是在所有裁判员给出的分数中去掉一个最低分和一个最高分，然后求其余分数的均值作为成绩．这一方法已被广泛应用，其原理可归为异常值理论，本文只对其中的期望估计量作一些探讨，给出几个重要性质．为叙述方便以下设 $EX=0$，非零情形只须作平移变换即可．

定义 1：设 $F(x)$ 为对称分布的分布函数，$X_1, X_2, \cdots, X_n$ iid$\sim F$．$X_{(1)}<X_{(2)}<\cdots<X_{(n)}$ 为顺序统计量，$F(x)$ 改进后的期望估计量：$W_p=\frac{1}{n-2[np]}\sum_{i=[np]+1}^{n-[np]}X_{(i)}$，其中 $0\leqslant p<\frac{1}{2}$。

1 W_p 的期望与方差

对一个新的估计量我们首先关心的是无偏性和方差大小，以下两个定理肯定了 W_p 的无偏性，并在 n，p，DX 确定的意义下给出了方差上界．

定理 1：设 $F(x)$ 为对称分布，且有密度函数 $f(x)$，$x_1, x_2, \cdots, x_n$ iid$\sim F$，则 $EW_p=0$ $(EX=0)$．

$$X_{(i)}\sim p_i(x_i)=\frac{n!}{(i-1)!(n-i)!}F^{i-1}(x_i)f(x_i)[1-F(x_i)]^{n-i}\qquad i=\overline{1,\ n}\quad \text{见[1]}$$

$$\therefore E(X_{(i)}+X_{(n-i+1)})=EX_{(i)}+EX_{(n-i+1)}$$

* 收稿日期：1989.10.21

$$=\int_{-\infty}^{+\infty}X_i\frac{n!}{(i-1)!(n-i)!}F^{i-1}(X_i)f(X_i)[1-F(X_i)]^{n-i}dX_i$$

$$+\int_{-\infty}^{+\infty}X_{n-i+1}\frac{n!}{(n-i)!(i-1)!}F^{n-i}(X_{n-i+1})f(X_{n-i+1})[1-F(X_{n-i+1})]^{i-1}dX_{n-i+1}$$

$$=\frac{n!}{(i-1)!(n-i)!}\int_{-\infty}^{+\infty}yf(y)\{F^{i-1}(y)[1-F(y)]^{n-i}+F^{n-i}(y)[1-F(y)^{i-1}\}dy$$

以上各积分的存在是显然的，注意到$f(y)$及$F^{i-1}(y)[1-F(y)]^{n-i}+F^{n-i}(y)\cdot[1-F(y)]^{i-1}$在$(-\infty,+\infty)$上是偶函数，故末式积分为0，即$E(X_{(i)}+X_{(n-i+1)})=0$。

所以：$$EW_p=E[\frac{1}{n-2[np]}\sum_{i=[np]+1}^{n-[np]}X_{(i)}]=E[\frac{1}{2(n-[np])}\sum_{i=[np]+1}^{n-[np]}(X_{(i)}+X_{(n-i+1)})]$$

$$=\frac{1}{2(n-[np])}\sum_{i=[np]+1}^{n-[np]}E(X_{(i)}+X_{(n-i+1)})=0$$

定理2：设F对称分布且有密度函数$f(x)$，X_1，X_2，…，X_n iid$\sim F$，$X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n)}$为其顺序统计量，$DX_1=\sigma^2$，则：

$$DW_p\leqslant\frac{n!}{([np]!)^2(n-2[np])!4^{[np]}(n-2[np])}\cdot\sigma^2$$

证：$EW_p=0$，设$[np]=r$。

$$(X_{(r+1)},X_{(r+2)},\cdots,X_{(n-r)})\sim p(X_{r+1},X_{r+2},\cdots,X_{n-r})$$

$$=\frac{n!}{r!r!}F^r(X_{r+1}f(X_{r+1})f(X_{r+2})\cdots f(X_{n-r})[1-F(X_{n-r})]^r$$

这里 $X_{r+1}<X_{r+2}<\cdots<X_{n-r}$ 见[1]

所以：$$DW_p=\frac{1}{(n-2r)^2}\int_{x_{r+1}<\cdots<x_{n-r}}(\sum_{i=r+1}^{n-r}X_i)^2\frac{n!}{r!r!}F^r(X_{r+1})f(X_{r+1})\cdots$$

$$f(X_{n-r})[1-F(X_{n-r})]^r dx_{r+1}dx_{r+2}\cdots dx_{n-r}$$

因 $0\leqslant F(X_{r+1})\leqslant F(X_{n-r})\leqslant 1$，$0\leqslant F(X_{r+1})[1-F(X_{n-r})]\leqslant\frac{1}{4}$

故 $$DW_p\leqslant\frac{n!}{(n-2r)^2(r!)^2 4^{2r}}\int_{x_{r+1}<\cdots<x_{n-r}}(\sum_{i=r+1}^{n-r}X_i)^2 f(X_{r+1})\cdots f(X_{n-r})dx_{r+1}\cdots dx_{n-r}$$

又因 $$R^{n-2r}=\sum_{i_{r+1}\cdots i_{n-r}}\{Xi_{r+1},Xi_{r+2},\cdots,Xi_{n-r})|Xi_{r+1}<Xi_{r+2}<\cdots<Xi_{n-r}\}$$

以上集的和是对$r+1$，$r+2$，…，$n-r$的全排列i_{r+1}，i_{r+2}，…，i_{n-r}求的，其中各集上积分是相等的，故有：

$$(n-2r)!\int_{x_{r+1}<\cdots<x_{n-r}}(\sum_{i=r+1}^{n-r}X_i)^2 f(X_{r+1})\cdots f(X_{n-r})dx_{r+1}\cdots dx_{n-r}$$

$$= \int_{R^{n-2r}} (\sum_{i=r+1}^{n-r} X_i)^2 f(X_{r+1}) \cdots f(X_{n-r}) dx_{r+1} \cdots dx_{n-r} = (n-2r)\sigma^2$$

所以　$DW_p \leq \frac{n!}{(r!)^2 (n-2r)! 4^{2r} (n-2r)} \sigma^2$，结论得证。

推论1：当$P=0$时，即$W_p = \overline{X}$，有$DW_p \leq \frac{\sigma^2}{n}$．　此推论与$D\overline{X} = \frac{\sigma^2}{n}$相符合．

2　W_p的强相合性

在大样本情况下估计量强相合性是非常重要的，以下定理在一定条件下肯定了W_p具有强相合性，其中的条件也较易验证。

引理：$P\{X_{(1)} < X_{(2)} < \cdots < X_{(n)}\} = 1$，$n = \overline{1,\ \infty}$．其中$F(x)$具有密度函数$f(x)$．

证：　$P\{X_{(1)} < X_{(2)} < \cdots < X_{(n)}\} = \int_{x_1 < x_2 < \cdots < x_n} n! f(X_1) f(X_2) \cdots f(X_n) dx_1 dx_2 \cdots dx_n$

$$= \int_{R^n} f(X_1) f(X_2) \cdots f(X_n) dx_1 \cdots dx_n = \prod_{i=1}^{n} \int_R f(X_i) dx_i = 1 \quad n < \infty \text{时得证.}$$

$$P\{X_{(1)} < X_{(2)} < \cdots < X_{(n)} < \cdots\} = P\{\bigcap_{n=1}^{\infty} (X_{(1)} < X_{(2)} < \cdots < X_{(n)})\}$$

$$= \lim_{n\to\infty} P\{X_{(1)} < X_{(2)} \cdots X_{(n)}\} = 1 \qquad n\to\infty \text{情形得证.}$$

此处用到了概率收敛定理．见[2]．

定理3：设$f(x)$严格单调增、连续且有密度函数$f(x)$，则W_p是EX的强相合估计．

证：(1)　$0 < p < \frac{1}{2}$．

由格列得科定理$P\{\lim_{n\to\infty} \sup_{x\in R} |F(x) - F_n(x)| = 0\} = 1$

又$F \in C$，↑，则F^{-1}在$[\frac{p}{2},\ 1-\frac{p}{2}]$上一致连续．所以有：

$$P\{\lim_{n\to\infty} \sup_{f(x),\ F_n(x)\in[\frac{p}{2},\ 1-\frac{p}{2}]} |X - F^{-1}(F_n(x))| = 0\} = 1$$

(2)　由$P\{\lim_{n\to\infty} |F(X_{([np]+1)}) - F_n(X_{([np]-1)})| = 0\} = 1$

$$F_n(X_{([np]-1)}) \stackrel{a.s.}{=\!=} \frac{[np]-2}{n} \to P, \quad n\to\infty \quad \text{(引理)}$$

所以　$P\{\lim_{n\to\infty} F(X_{([np]-1)}) = P\} = 1$

即：$F(X_{([np]-1)}) \xrightarrow{a.s} P$，　$F_n(X_{([np]-1)}) \xrightarrow{a.s} p$

同理：$F(X_{([(1-p)n]+1)}) \xrightarrow{a.s} 1-P$，　$F_n(X_{([(1-p)n]+1)}) \xrightarrow{a.s} 1-P$

由F(x)，$F_n(x)$的单调性知：$F(X_{(i)})$，$F_n(X_{(i)}) \in [\frac{P}{2}, 1-\frac{P}{2}]$　a.s.

$$i=\overline{[np]-1,\ [(1-p)n]+1}.$$

所以　$P\{\lim\limits_{n\to\infty} \sup\limits_{i=\overline{[np]-1,\ [(1-p)n]+1}} |X_{(i)}-F^{-1}(F_n(X_{(i)}))|=0\}=1$

(3)由引理知$F_n(X_{(i)})+F_n(X_{(n-i+1)})=Y_n$，　a.s.　又F(x)为对称分布.

故$F^{-1}(F_n(X_{(i)}))=-F^{-1}(F_n(X_{(n-i+1)}))+\triangle(\frac{1}{n})$，　a.s.

$$i=\overline{[np]-1,\ [(1-p)n]+1},\quad \triangle(\frac{1}{n})\to 0(n\to\infty)$$

故$|X_{(i)}+\ {}_{(n-i+1)}|\leqslant|X_{(i)}-F^{-1}(F_n(X_{(i)}))|+|X_{(n-i+1)}-F^{-1}(F_n(X_{(n-i+1)}))|+\triangle(\frac{1}{n})$

所以$P\{\lim\limits_{n\to\infty} \sup\limits_{i=\overline{[np]-1,\ [(1-p)n]+1}} |X_{(i)}+X_{(n-i+1)}|=0\}=1$

又$|W_p|=|\frac{1}{n-2[np]}\sum\limits_{i=[np]+1}^{n-[np]} X_{(i)}|=\frac{1}{2(n-[np])}|\sum\limits_{i=[np]+1}^{n-[np]}(X_{(i)}+X_{(n-i+1)})|$

$$\leqslant\frac{1}{2(n-[np])}|\sum_{i=[np]+1}^{n-[np]}(X_{(i)}+X_{(n-i+1)})|$$

所以$P\{|W_p| \xrightarrow{n\to\infty} 0\}=1$　即$W_p \xrightarrow{a.s} 0$．此处用到了$n-[np]\leqslant[(1-p)n]+1$

参 考 文 献

〔1〕华东师范大学编．数理统计讲义

〔2〕严士健，王隽骧，刘秀芳著．概率论基础．科学出版社

〔3〕朱宏．用样本分位方法同时检验正态样本多个异常值．数理统计与应用概率，1989年，第四卷，第一期，P30—40

Character of Modified Symmetry Distribution Expectation Estimation

Zhang Xinyu
(Zhengzhou Institute of Technology)

Abstract: This paper prove that the modified symmetry distribation expectation estimation is unbiased and strongly consistent for the condition so-and-so. It gives the upper bound of the new estimation Variance.

Keywords: mathematical expectation, estimations, unbiasedness