Journal of Statistical Planning and Inference 140 (2010) 2943–2951

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Review

Optimal sign test for quantiles in ranked set samples

Xiaofang Dong [a,b], Lirong Cui [a,*]

[a] School of Management and Economics, Beijing Institute of Technology, Beijing100081, China
[b] School of Sciences, Qiqihar University, China

ARTICLE INFO

Article history:
Received 18 March 2010
Received in revised form
10 May 2010
Accepted 12 May 2010
Available online 21 May 2010

Keywords:
Ranked set sampling
Nonparametric test
Quantile
Confidence interval
Pitman efficiency

ABSTRACT

This paper considers the one-sample sign test for population quantiles in general ranked set sampling, and proposes a weighted sign test because observations with different ranks are not identically distributed. It is shown analytically that optimal weight always improves the Pitman efficiency for all distributions. For each quantile, the sampling allocation that maximizes the sign test efficacy is identified and shown to not depend on the population distribution. Moreover, distribution-free confidence intervals for quantiles based on ordered values of optimal ranked set samples are discussed.

Contents

1. Introduction

Ranked set sampling (RSS) is a sampling protocol that can often be used to improve the cost efficiency of an experiment. It is appropriate for situations in which quantification of sampling units is costly or difficult but ranking of the units in a small set is easy and inexpensive.

General RSS (unequal allocation) involves selecting $n \times m$ units at random from an infinite population. These units are divided into n sets, each having m units. In each set, units are ranked from smallest to largest by some auxiliary criterion that does not require actual measurements. From within each of n sets, exactly one unit is quantified and this unit is chosen according to its rank-order. The number of times rank-order i is to be quantified is n_i ($\sum_{i=1}^{m} n_i = n$). Thus, we obtain

* Corresponding author. Tel.: +86 0 10 68940780.
E-mail address: Lirongcui@bit.edu.cn (L. Cui).

0378-3758/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2010.05.005

ranked set sample $X_{(i)j}, j=1,\ldots,n_i;\ i=1,\ldots,m$. The standard RSS (equal allocation) is a special case when $n_1=n_2=\cdots=n_m$. If the judgment ranking is perfect, the ith judgment order statistic is the same as the ith order statistic having m units. The set-size, m, is generally small for implementation reason, including ease of ranking.

The idea of RSS was first proposed by McIntyre (1952) in order to find a more efficient method to estimate the average yield of pasture. Since then, many statistical procedures including parametric and nonparametric procedures based on the RSS have been investigated in the literature by many others, namely, Takahasi and Wakimoto (1968), Dell and Clutter (1972), Stokes (1980, 1995), Stokes and Sager (1988), Bohn and Wolfe (1992, 1994), Chen (1999), and Balakrishnan and Li (2005, 2008).

Recently, the sign test, one of the fundamental methods in nonparametric, has been studied for analyzing RSS data by some authors. Hettmansperger (1995) first considered the sign test for the standard RSS data and found that the Pitman efficiency of the sign test can be improved substantially. Ozturk (1999) considered the sign test in selective designs, which selects only a subset of the ranks among 1 to m but the number of observations are assumed to be the same for the selected ranks, general RSS includes the selective designs as special cases. Ozturk and Wolfe (2000) and Wang and Zhu (2005) showed that the Pitman efficiency achieves the highest when selecting the median observations for quantification. Above results all were about the sign test for testing the median of a population, but quantile testing has frequent applications in practical assessments. For example, making inference on a population median may be an attractive alternative to making inference on a population mean when the underlying distribution is highly skewed, and making inference on extreme quantiles of a distribution is a way to assess the prevalence of dangerously high or low values in a population. Chen (2000) discussed ranked-set sample quantiles and their applications. Balakrishnan and Li (2006) considered distribution-free confidence intervals for quantiles based on ordered RSS. Ozturk and Balakrishnan (2009) proposed an exact two-sample nonparametric test for quantiles under the standard RSS.

In this paper, we will consider the one sample sign test for population quantiles based on general RSS. In Section 2, we provide analytical proof that Pitman efficiency of the sign test for RSS data can be improved by weighting the observations based on their associated ranks. For each quantile, the optimal sampling allocation that maximizes the sign test efficacy is identified and shown to not depend on the population distribution in Section 3. Section 4 focuses on the confidence intervals for quantiles based on ordered values of the optimal RSS and their properties.

2. The sign test for population quantiles

Let ξ_p be a known constant, the null hypothesis asserts that ξ_p is the pth quantile of infinite population having cumulative distribution function (*cdf*) $F(x)$ and probability density function (*pdf*) $f(x)$. Thus, we want to test $H_0{:}F(\xi_p)=p$ against either a one-sided or a two-sided alternative. For $p=0.5$, this scenario reduces to a test of the median.

Let $X_1,X_2,\ldots,X_n$ be a simple random sample (SRS) drawn from the continuous distribution $F(x)$, the sign test statistic under SRS is given by

$$S_{SRS}^{+}=\sum_{i=1}^{n}\psi(X_i-\xi_p),$$

where $\psi(t)$ is the indicator function $I(t>0)$.

It is easy to verify that $S_{SRS}^{+}\sim binomial(n,1-F(\xi_p))$ and the Pitman efficacy of S_{SRS}^{+} is

$$eff(S_{SRS}^{+})=\frac{f^2(\xi_p)}{p(1-p)}\Big|_{H_0},$$

where the efficacy $eff(T)$ of a test statistic T is defined by

$$eff(T)=\frac{[\partial E(T)/\partial\xi_p]^2}{nVar(T)}\Big|_{H_0}. \tag{2.1}$$

Let $X_{(i)j}, j=1,\ldots,n_i$ and $i=1,\ldots,m$ be general RSS of size $n=\sum_{i=1}^{m}n_i$. We assume perfect ranking, so that $X_{(i)j}$ is the jth observation on ith order statistic of F. The distribution of $X_{(i)j}$, which depends on the rank order i but not on j, has its *pdf* and *cdf* as follows:

$$f_{(i)}(x)=\frac{1}{B(i,m+1-i)}[F(x)]^{i-1}[1-F(x)]^{m-i}f(x), \tag{2.2}$$

$$F_{(i)}(x)=\int_{-\infty}^{x}f_{(i)}(t)\,dt=I_{F(x)}(i,m+1-i), \tag{2.3}$$

where $B(a,b)=\int_0^1 t^{a-1}(1-t)^{b-1}\,dt$ is complete beta function and $I_u(a,b)$, called as incomplete beta function, is defined as

$$I_u(a,b)=\frac{1}{B(a,b)}\int_0^u t^{a-1}(1-t)^{b-1}\,dt,\quad 0\le u\le 1. \tag{2.4}$$