Ann Inst Stat Math (2011) 63:227–243
DOI 10.1007/s10463-008-0215-z

Asymptotic properties of sample quantiles of discrete distributions

Yanyuan Ma · Marc G. Genton · Emanuel Parzen

Received: 23 May 2008 / Revised: 14 October 2008 / Published online: 22 January 2009

Abstract The asymptotic distribution of sample quantiles in the classical definition is well-known to be normal for absolutely continuous distributions. However, this is no longer true for discrete distributions or samples with ties. We show that the definition of sample quantiles based on mid-distribution functions resolves this issue and provides a unified framework for asymptotic properties of sample quantiles from absolutely continuous and from discrete distributions. We demonstrate that the same asymptotic normal distribution result as for the classical sample quantiles holds at differentiable points, whereas a more general form arises for distributions whose cumulative distribution function has only one-sided differentiability. For discrete distributions with finite support, the same type of asymptotics holds and the sample quantiles based on mid-distribution functions either follow a normal or a two-piece normal distribution. We also calculate the exact distribution of these sample quantiles for the binomial and Poisson distributions. We illustrate the asymptotic results with simulations.

Keywords Asymptotic · Continuous · Discrete · Grouped data · Mid-distribution · Quantile · Ties

Y. Ma · M. G. Genton (✉) · E. Parzen
Department of Statistics, Texas A&M University, College Station, TX 77843, USA
e-mail: genton@stat.tamu.edu

Y. Ma
e-mail: ma@stat.tamu.edu

E. Parzen
e-mail: eparzen@stat.tamu.edu

1 Introduction

Sample quantiles were pioneered by Galton (1889). According to Hald (1998), Sheppard in 1890 may have been the first to state the asymptotic distribution of sample quantiles of a continuous random variable. The classical definition of sample quantiles and their asymptotic properties for continuous distributions are nowadays well-known in the literature; see, e.g., David and Nagaraja (2003) and further references therein. However, many drawbacks occur with the classical definition, mainly in the case of tied samples and discrete distributions. For illustration, consider a binary sample of size five: 0, 0, 0, 0, 1. The classical definition yields a sample median of 0. The same answer is obtained for the sample 0, 0, 0, 1, 1. However, imagine such patterns repeat and the sample size becomes much larger. A more rational answer should be 1/5 and 2/5, respectively, the empirical proportion of 1's in the sample, in order to reflect the difference between the two samples and their underlying distributions. Such an answer can indeed be obtained by using an alternative definition of sample median based on mid-distribution functions (Parzen 1997, 2004), a tool inspired by the concept of mid-p-value introduced by Lancaster (1961); see also Berry and Armitage (1995) for a review. The large sample behavior of such a sample median was observed to be close to normal in some numerical examples in Genton et al. (2006). They also showed by means of Monte Carlo simulations that on the contrary, the asymptotic distribution of the classical sample median is not of normal type, but a discrete distribution.

The goal of our paper is to establish the asymptotic properties of sample quantiles based on mid-distribution functions, for both continuous and discrete distributions. We show that for an absolutely continuous distribution function and any quantiles where the distribution function is differentiable, the same asymptotic property holds for these sample quantiles as for the classical sample quantiles. This asymptotic property is shown to hold in a more relaxed form if the distribution function is only one-sided differentiable at these quantiles. However, for the quantiles where the distribution function is discrete with finite support, these sample quantiles behave more favorably than the classical sample quantiles. That is, while no asymptotic normality properties are available for the classical sample quantiles, such properties are shown to hold in various forms for sample quantiles based on mid-distribution functions. Many studies are based on grouped data, hence implying discrete data from an underlying continuous distribution. This is important, for example, when constructing confidence intervals for the median of a discrete population, see Larocque and Randles (2008). The issue of computing quantiles from samples with ties also emerges in the study of copulas for count data, see Genest and Nešlehová (2007). Note that González-Barrios and Rueda (2001) have studied convergence theorems for quantiles in the case they are not unique. However, in the discrete case, their approach does not result in the natural solution we advocate above. Also Machado and Santos Silva (2005) have considered quantiles for counts by artificially imposing smoothness in the problem through jittering followed by kernel interpolation of the quantile function. However, their approach is not based on quantiles defined through mid-distribution functions.

The rest of the paper is organized as the following. In Sect. 2, we explain the definition of sample quantiles based on mid-distribution functions and define theoretical quantiles under various assumptions on the underlying distribution. We state and prove the asymptotic properties of these sample quantiles for absolutely continuous distributions when the cumulative distribution function is differentiable or one-sided differentiable in Sect. 3. Although similar results can be found in Smirnov (1952), our formulation is modern and given for the sake of completeness. We establish new asymptotic properties when the distribution is discrete with finite support in Sect. 4. The exact distribution of the mid-distribution based sample quantiles for binomial and Poisson distributions is given in Sect. 5. Section 6 concentrates on a simulation study to demonstrate the relevance of these asymptotic results.

2 Sample quantiles based on mid-distribution functions

We consider first a definition of population quantiles based on the concept of mid-distribution function. Let $F(x) = P(X \leq x)$ be the cumulative distribution function of a random variable X with probability mass function $p(x) = P(X = x)$ and probability density function $f(x) = F'(x)$ for differentiable F. Note that X can be either continuous or discrete. Define the mid-distribution function $F_{\text{mid}}(x) = F(x) - 0.5p(x)$. When the random variable X is continuous, $p(x) \equiv 0$ and thus $F_{\text{mid}}(x) = F(x)$. In that case, the quantile function is $Q(p) = F^{-1}(p)$, $0 \leq p \leq 1$, for strictly monotone F, and $Q(p) = \inf_t\{F(t) \geq p\}$ for monotone F. When the random variable X is discrete, let $v_1 < \cdots < v_d$ denote its distinct values with corresponding probabilities of occurrence $p_1, \ldots, p_d$. In that case, the quantile function $Q(p)$ is defined as being piecewise linear and connecting the values $Q(F_{\text{mid}}(v_j)) = v_j$, $j = 1, \ldots, d$. Specifically, we define the quantile function for a discrete distribution by

$$Q(p) = F_{\text{mid}}^{-1}(p) = \begin{cases} v_1 & \text{if } p < p_1/2, \\ v_k & \text{if } p = \pi_k, k = 1, \ldots, d, \\ \lambda v_k + (1-\lambda)v_{k+1} & \text{if } p = \lambda\pi_k + (1-\lambda)\pi_{k+1}, 0 < \lambda < 1, \\ & \quad k = 1, \ldots, d-1, \\ v_d & \text{if } p > \pi_d, \end{cases}$$

where $\pi_k = \sum_{i=1}^{k-1} p_i + p_k/2$, $k = 1, \ldots, d$, are called the mid-p-values. Note that $d = \infty$ is allowed, in which case the last category will be suppressed (see Fig. 1 for an illustration).

In the sample version, the sample cumulative distribution function of $X_1, \ldots, X_n$ is defined as $\tilde{F}(x) = \tilde{P}(X \leq x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x)$ where $I(\cdot)$ denotes the indicator function. Here, we use $\tilde{}$ to denote sample version empirical calculations. Similarly, the sample probability mass function is $\tilde{p}(x) = \tilde{P}(X = x)$ and the sample mid-distribution function is $\tilde{F}_{\text{mid}}(x) = \tilde{F}(x) - 0.5\tilde{p}(x)$. Note that there is no corresponding sample probability density function. For samples with distinct values, Parzen (2004) defines the sample quantile function $\tilde{Q}(p)$ as being piecewise linear and connecting the values $\tilde{Q}\{(j-0.5)/n\} = X_{j:n}$, where $X_{1:n} \leq \cdots \leq X_{n:n}$ are the order statistics of the sample, i.e.,

$$\tilde{Q}(p) = (0.5 - np + j)X_{j:n} + (0.5 + np - j)X_{j+1:n}, \tag{1}$$

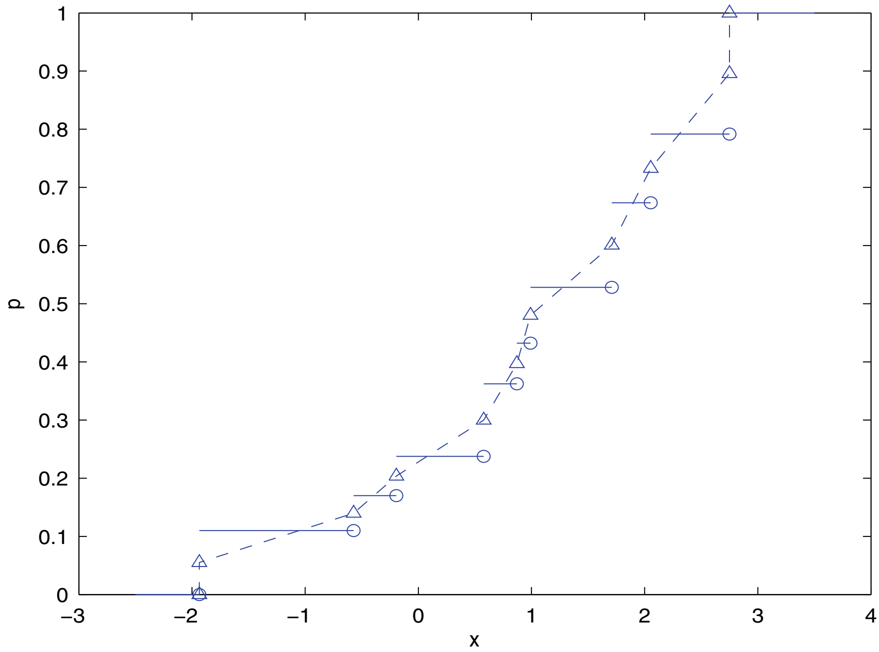

Fig. 1 An illustration of the mid-distribution function. The *solid horizontal lines* form the cumulative distribution function, where the *circles* are the jumps. The *dashed line* illustrates the mid-distribution function, where the *triangles* indicate the values of the mid-distribution function at the supporting points

for $(j-0.5)/n < p \leq (j+0.5)/n$. For samples with ties, denote the distinct values in the sample by $v_1 < \cdots < v_d$ with corresponding frequencies $r_1, \ldots, r_d$. Then, define the sample quantile function $\tilde{Q}(p)$ as being piecewise linear and connecting the values $\tilde{Q}\{\tilde{F}_{\text{mid}}(v_j)\} = v_j$, $j = 1, \ldots, d$, which can be viewed as a definition of fractional order statistics. Although it is a bit tedious to write down the analytic form of $\tilde{Q}$, the concept is very simple to grasp. Specifically, let $v_1 < \cdots < v_d$ be realizations of the random variables $V_1 < \cdots < V_d$. Then $\tilde{Q}$ is calculated as the following:

- *Step 1* Calculate the sample cumulative distribution function $\tilde{F}(v_1) = r_1/n, \ldots, \tilde{F}(v_k) = \sum_{i=1}^{k} r_i/n, \ldots, \tilde{F}(v_d) = 1$.
- *Step 2* Calculate the sample probability mass function $\tilde{p}(v_1) = r_1/n, \ldots, \tilde{p}(v_k) = r_k/n, \ldots, \tilde{p}(v_d) = r_d/n$.
- *Step 3* Calculate the sample mid-distribution function $\tilde{F}_{\text{mid}}(v_1) = r_1/(2n), \ldots, \tilde{F}_{\text{mid}}(v_k) = \sum_{i=1}^{k} r_i/n - r_k/(2n), \ldots, \tilde{F}_{\text{mid}}(v_d) = 1 - r_d/(2n)$.
- *Step 4* Find $k, 0 \leq k \leq d$, such that $p = \lambda \tilde{F}_{\text{mid}}(v_k) + (1-\lambda)\tilde{F}_{\text{mid}}(v_{k+1})$ for $0 \leq \lambda < 1$. Here, we set $v_0 = v_1$ and $v_{d+1} = v_d$.
- *Step 5* Calculate the pth sample quantile as $\tilde{Q}(p) = \lambda v_k + (1-\lambda)v_{k+1}$.

From the above definition, for $p < r_1/(2n)$ or $p > 1 - r_d/(2n)$, the corresponding sample quantile is defined as v_1 or v_d.

The classical sample quantile function, denoted by $\tilde{Q}(p) = \tilde{F}^{-1}(p)$, is defined to be a piecewise constant function such that $\tilde{Q}(p) = X_{j:n}$ for $(j-1)/n < p \leq j/n$.

Unfortunately, the sample median $\tilde{Q}(0.5)$ based on this definition does not agree with the usual definition: $X_{m+1:n}$ if $n = 2m + 1$ and $0.5(X_{m:n} + X_{m+1:n})$ if $n = 2m$. However, the definition (1) does agree with the usual definition of the sample median. More importantly, Genton et al. (2006) have suggested that the sample median based on the classical definition above is inadequate in handling samples with ties, which typically occur from discrete distributions or grouped data. By means of various simulation experiments, they have illustrated that the classical asymptotic results do not hold in the setting of ties. In particular, the limiting distribution in this case appears to be discrete. Unlike the classical definition, the proposal for sample quantiles based on mid-distribution functions is designed to handle samples with ties, and we show in Sect. 4 that in most cases, its limiting distribution is continuous and the estimator has the typical root-n consistency and asymptotic normality.

One can certainly search for many alternative ways to estimate the mid-distribution based quantiles, instead of using the sample version definition. However, since the population quantile definition based on mid-distribution is different from the classical one, any familiar existing method would have to be carefully modified. The sample quantile described above is simple and natural, hence this is what we recommend.

3 Absolutely continuous case

In this section, we demonstrate that in the classical case, where near a certain point ξ the cumulative distribution function F is differentiable and the density f is positive, the sample quantiles based on mid-distribution functions have the same asymptotic properties as the classical sample quantiles. We also show that when F is left-differentiable and right-differentiable, a different form of the same asymptotic result holds. The proof mainly follows the lines of Walker (1968).

Theorem 1 *Assume that $X_1, \ldots, X_n$ is a random sample from a distribution with absolutely continuous cumulative distribution F. Denote the pth sample quantile based on mid-distribution functions by $\hat{\xi}$.*

Case 1 *Assume that in a neighborhood of ξ, $F(x)$ is differentiable and $f(x) = F'(x) > 0$. Let $p = F(\xi)$. Then, we have*

$$\sqrt{n}(\hat{\xi} - \xi) f(\xi) \to N\{0, p(1-p)\}$$

in distribution when $n \to \infty$.

Case 2 *Assume that $F(x)$ has one-sided derivatives $f_+(\xi) = \lim_{x \to \xi^+}\{F(x) - F(\xi)\}/(x-\xi) > 0$ and $f_-(\xi) = \lim_{x\to\xi^-}\{F(x) - F(\xi)\}/(x-\xi) > 0$. In addition, assume that in a neighborhood of ξ, $F(x)$ is differentiable and $f(x) = F'(x) > 0$ for $x \neq \xi$. Let $p = F(\xi)$. Then*

$$\sqrt{n}(\hat{\xi} - \xi) f(\hat{\xi}, \xi) \to N\{0, p(1-p)\}$$

in distribution when $n \to \infty$. Here, we denote $f(\hat{\xi}, \xi) = f_+(\xi)$ if $\hat{\xi} > \xi$ and $f(\hat{\xi}, \xi) = f_-(\xi)$ if $\hat{\xi} < \xi$.

Proof

Case 1 Using the notation from Sect. 2, assume

$$p = \lambda_n \tilde{F}_{\text{mid}}(V_k) + (1-\lambda_n)\tilde{F}_{\text{mid}}(V_{k+1}) \\ = \lambda_n \left(\frac{r_1 + \cdots + r_k}{n} - \frac{r_k}{2n} \right) + (1-\lambda_n)\left(\frac{r_1 + \cdots + r_k}{n} + \frac{r_{k+1}}{2n} \right), \tag{2}$$

hence $\hat{\xi} = \lambda_n V_k + (1-\lambda_n)V_{k+1}$. For an arbitrary fixed quantity z, we have

$$\Pr(\hat{\xi} < \xi + n^{-1/2}z) = \Pr\{F(\hat{\xi}) < F(\xi + n^{-1/2}z)\}. \tag{3}$$

Since $V_k \leq \hat{\xi} \leq V_{k+1}$, so $F(V_k) \leq F(\hat{\xi}) \leq F(V_{k+1})$. Therefore we have

$$\Pr\{F(V_{k+1}) < F(\xi + n^{-1/2}z)\} \leq \Pr\{F(\hat{\xi}) < F(\xi + n^{-1/2}z)\} \\ \leq \Pr\{F(V_k) < F(\xi + n^{-1/2}z)\}.$$

Next, we prove that $\Pr\{F(V_j) < F(\xi + n^{-1/2}z)\} \to \Phi\{zf(\xi)/\sqrt{p(1-p)}\}$ when $n \to \infty$ for both $j = k$ and $j = k+1$.

Note that $F(X_i)$, $i = 1, \ldots, n$, is a random sample from a uniform distribution on [0, 1]. Let N be the number of $F(X_i)$'s that are smaller than $F(\xi + n^{-1/2}z)$. Then N follows a binomial distribution with parameters n and $F(\xi + n^{-1/2}z)$. As $n \to \infty$, denoting $c = F(\xi + n^{-1/2}z)$, standard results (see for example, Parzen 1954) imply that

$$\frac{N - nc}{\sqrt{nc(1-c)}} \to N(0, 1)$$

in distribution uniformly in c. Thus we have

$$\Pr\{F(V_j) < c\} = \Pr\{N \geq (r_1 + \cdots + r_j)\} \\ = \Pr\left\{ \frac{N - nc}{\sqrt{nc(1-c)}} \geq \frac{r_1 + \cdots + r_j - nc}{\sqrt{nc(1-c)}} \right\} \\ = \Phi\left\{ \frac{nc - (r_1 + \cdots + r_j)}{\sqrt{nc(1-c)}} \right\} + o_p(1)$$

as $n \to \infty$. Since $c = p + n^{-1/2}z\{f(\xi) + \epsilon\}$, where $\epsilon \to 0$ as $n \to \infty$, we have

$$\Pr\{F(V_j) < c\} = \Phi\left\{ \frac{np - (r_1 + \cdots + r_j)}{\sqrt{n}\sqrt{p(1-p)}} + \frac{zf(\xi)}{\sqrt{p(1-p)}} \right\} + o_p(1).$$

It is easily verified that $|np - (r_1 + \cdots + r_j)| < r_k + r_{k+1}$ for both $j = k$ and $j = k+1$. In addition, because $F(x)$ is differentiable and $f(x)$ is positive in the neighborhood of ξ, $r_k = 1$ and $r_{k+1} = 1$ with probability 1. Thus, for sufficiently large n, $\{np - (r_1 + \cdots + r_j)\}/\sqrt{n} = o_p(1)$. Consequently,

$$\Pr\{F(V_j) < c\} \to \Phi\left\{\frac{zf(\xi)}{\sqrt{p(1-p)}}\right\}$$

for $j = k$ and $j = k + 1$. It now follows that

$$\Pr(\hat{\xi} < \xi + n^{-1/2}z) = \Pr\{F(\hat{\xi}) < F(\xi + n^{-1/2}z)\} \to \Phi\left\{\frac{zf(\xi)}{\sqrt{p(1-p)}}\right\}$$

and thus completes the proof.

Case 2 Similar to case 1, we assume (2), hence $\hat{\xi} = \lambda_n V_k + (1 - \lambda_n)V_{k+1}$. For an arbitrary fixed quantity z, we consider separately two cases: $z \leq 0$ and $z > 0$.

When $z \leq 0$, we have (3). The proof then follows the steps of case 1 with $f(\xi)$ replaced by $f_-(\xi)$ and noting that, because $f_+(\xi)$, $f_-(\xi)$ and $f(x)$ are positive in the neighborhood of ξ, $r_k = 1$ and $r_{k+1} = 1$ with probability 1. Thus, we obtain, as $n \to \infty$,

$$\Pr\{F(V_j) < c\} \to \Phi\left\{\frac{zf_-(\xi)}{\sqrt{p(1-p)}}\right\}$$

for $j = k$ and $j = k + 1$. It now follows that

$$\begin{aligned}\Pr\{f(\hat{\xi}, \xi)\hat{\xi} < f(\hat{\xi}, \xi)\xi + n^{-1/2}zf_-(\xi)\} &= \Pr(\hat{\xi} < \xi + n^{-1/2}z)\\ &= \Pr\{F(\hat{\xi}) < F(\xi + n^{-1/2}z)\}\\ &\to \Phi\left\{\frac{zf_-(\xi)}{\sqrt{p(1-p)}}\right\},\end{aligned}$$

or equivalently,

$$\Pr\left\{\frac{\sqrt{n}f(\hat{\xi}, \xi)(\hat{\xi} - \xi)}{\sqrt{p(1-p)}} < z\right\} \to \Phi(z)$$

where $n \to \infty$ for $z \leq 0$.

When $z > 0$, we have

$$\begin{aligned}&\Pr\left\{\frac{\sqrt{n}f(\hat{\xi}, \xi)(\hat{\xi} - \xi)}{\sqrt{p(1-p)}} < z\right\}\\ &\quad = \Pr\left\{\frac{\sqrt{n}f(\hat{\xi}, \xi)(\hat{\xi} - \xi)}{\sqrt{p(1-p)}} < 0\right\} + \Pr\left\{0 < \frac{\sqrt{n}f(\hat{\xi}, \xi)(\hat{\xi} - \xi)}{\sqrt{p(1-p)}} < z\right\}\\ &\quad = \Pr\left\{\frac{\sqrt{n}f_-(\xi)(\hat{\xi} - \xi)}{\sqrt{p(1-p)}} < 0\right\} + \Pr\left\{0 < \frac{\sqrt{n}f_+(\xi)(\hat{\xi} - \xi)}{\sqrt{p(1-p)}} < z\right\}\end{aligned}$$

$$= \Pr\left\{\frac{\sqrt{n}f_+(\xi)(\hat{\xi}-\xi)}{\sqrt{p(1-p)}} < 0\right\} + \Pr\left\{0 < \frac{\sqrt{n}f_+(\xi)(\hat{\xi}-\xi)}{\sqrt{p(1-p)}} < z\right\}$$
$$= \Pr\left\{\frac{\sqrt{n}f_+(\xi)(\hat{\xi}-\xi)}{\sqrt{p(1-p)}} < z\right\} \to \Phi(z)$$

in distribution when $n \to \infty$, where the last step can be derived by following exactly the same procedure as for the case of $z \leq 0$, and replacing $f_-(\xi)$ by $f_+(\xi)$ everywhere. □

Remark 1 The result in case 1 of Theorem 1 can be alternatively expressed as

$$\sqrt{n}(\hat{\xi}-\xi) \to N\left\{0, \frac{p(1-p)}{f^2(\xi)}\right\}$$

in distribution when $n \to \infty$, which is the more familiar form in the literature.

There are also other boundary cases to consider. For example, what kind of asymptotic results would hold if $f_-(\xi) = 0$ or $f_+(\xi) = 0$? What happens if $f(x) > 0$ does not hold for an arbitrary neighborhood of ξ? Instead of digging into all these cases, we study a very important extreme case where the distribution is discrete with finite support. Thus, the various boundary cases can be considered as a combination of the different scenarios that we study.

4 Discrete case

When the true distribution is discrete, it has been observed that the classical sample median does not follow the typical asymptotic normal distribution. In fact, it is not clear that a systematic asymptotic behavior of the classical sample quantiles is present. However, the situation is more optimistic for sample quantiles based on mid-distribution functions and we present their asymptotic distributions in Theorem 2. In order to express the results in the same form as in Theorem 1, we first introduce some notations.

We define $f(\xi) \equiv F'_{\text{mid}}(\xi) \equiv \frac{1}{2}(p_k + p_{k+1})/(v_{k+1} - v_k)$ for an arbitrary ξ satisfying $v_k < \xi < v_{k+1}$, $f_+(v_k) \equiv F'_{\text{mid}+}(v_k) \equiv \frac{1}{2}(p_k + p_{k+1})/(v_{k+1} - v_k)$, $f_-(v_k) \equiv F'_{\text{mid}-}(v_k) \equiv \frac{1}{2}(p_k + p_{k-1})/(v_k - v_{k-1})$ and we define $f(\hat{\xi}, v_k) = f_+(v_k)$ if $\hat{\xi} > v_k$ and $f(\hat{\xi}, v_k) = f_-(v_k)$ if $\hat{\xi} < v_k$. Recall that $\pi_k = \sum_{i=1}^{k-1} p_i + p_k/2$, $k = 1, \ldots, d$, are the mid-p-values.

Theorem 2 *Assume we generate a sample of n independent observations from a discrete distribution. The support of the distribution is $v_1 < \cdots < v_d$, and the corresponding probabilities are $p_1, \ldots, p_d$. Define $p_0 = 0$, $p_{d+1} = 0$, $r_0 = 0$, $r_{d+1} = 0$, $v_0 = v_1$, $v_{d+1} = v_d$. Then when $n \to \infty$, the pth sample quantile based on mid-distribution functions $\hat{\xi}$ will*

Case 1 *Converge to v_1 in probability if $p < p_1/2$.*
Case 2 *Converge to v_d in probability if $p > \pi_d$.*

Case 3 *Satisfy*

$$\sqrt{n}(\hat{\xi}-\xi)f(\xi) \to N\Big[0,\, p(1-p)-\{1-(\lambda-1)^2\}p_{k+1}/4 \\ -(1-\lambda^2)p_{k+2}/4\Big]$$

in distribution as $n \to \infty$ *if*

$$p = \lambda\pi_{k+1} + (1-\lambda)\pi_{k+2}$$

for $0 < \lambda < 1$, $k = 0, \ldots, d-2$. *Here,* $\xi = \lambda v_{k+1} + (1-\lambda)v_{k+2}$.

Case 4 *Satisfy* $\sqrt{n}(\hat{\xi}-v_{k+1})f(\hat{\xi}, v_{k+1}) \to N(0, p(1-p)-p_{k+1}/4)$ *in distribution as* $n \to \infty$ *if*

$$p = \pi_{k+1}, \quad \text{for } k = 0, \ldots, d-1.$$

Proof

Case 1 When n is sufficiently large, with probability tending to 1, we can find $0 < \lambda_n < 1$ such that $p = \lambda_n r_0/n + (1-\lambda_n)r_1/(2n)$. Hence with probability tending to 1, $\hat{\xi} = \lambda_n v_0 + (1-\lambda_n)v_1 = v_1$.

Case 2 When n is sufficiently large, with probability tending to 1, we can find $0 < \lambda_n < 1$ such that $p = \lambda_n\{(r_1 + \cdots + r_{d-1})/n + r_d/(2n)\} + (1-\lambda_n)\{(r_1 + \cdots + r_d)/n + r_{d+1}/(2n)\}$. Hence with probability tending to 1, $\hat{\xi} = \lambda_n v_d + (1-\lambda_n)v_{d+1} = v_d$.

Case 3 Denote $s_i^2 = p_i(1-p_i)$, $i = 1, \ldots, d$. When n is sufficiently large, with probability tending to 1, we can find $0 < \lambda_n < 1$ such that

$$p = \lambda_n\left(\frac{r_1 + \cdots + r_k}{n} + \frac{r_{k+1}}{2n}\right) + (1-\lambda_n)\left(\frac{r_1 + \cdots + r_{k+1}}{n} + \frac{r_{k+2}}{2n}\right),$$

thus

$$\lambda_n = \left(\frac{r_1 + \cdots + r_{k+1}}{n} + \frac{r_{k+2}}{2n} - p\right)\Big/\left(\frac{r_{k+1} + r_{k+2}}{2n}\right) \quad \text{and}$$
$$\hat{\xi} = \lambda_n v_{k+1} + (1-\lambda_n)v_{k+2}.$$

Since $\hat{\xi} - \xi = \hat{\xi} - \lambda v_{k+1} - (1-\lambda)v_{k+2} = (\lambda_n - \lambda)(v_{k+1} - v_{k+2})$, we only need to investigate

$$\lambda_n - \lambda = \left(\frac{r_1 + \cdots + r_{k+1}}{n} + \frac{r_{k+2}}{2n} - p\right)\Big/\left(\frac{r_{k+1} + r_{k+2}}{2n}\right) \\ -\left(p_1 + \cdots + p_{k+1} + \frac{p_{k+2}}{2} - p\right)\Big/\left(\frac{p_{k+1} + p_{k+2}}{2}\right).$$

Clearly, $r_1, \dots, r_{k+2}$ is a random sample from a multinomial distribution with means $np_1, \dots, np_{k+2}$, variances $np_1(1-p_1), \dots, np_{k+2}(1-p_{k+2})$ and covariances $-np_i p_j$ for $i \neq j, i, j = 1, \dots, k+2$. Hence when $n \to \infty$, $r_1/n, \dots, r_{k+2}/n$ converges in distribution to a multivariate normal distribution with mean vector $(p_1, \dots, p_{k+2})$, variances $p_1(1-p_1), \dots, p_{k+2}(1-p_{k+2})$ and covariances $-p_i p_j$ for $i \neq j, i, j = 1, \dots, k+2$. Using the delta method, through straightforward calculation, we immediately obtain

$$\sqrt{n}(\lambda_n - \lambda) \to N(0, \sigma^2)$$

where

$$\sigma^2 = (p_{k+1} + p_{k+2})^{-2} \left\{4p(1-p) + \lambda(\lambda - 2)p_{k+1} + (\lambda^2 - 1)p_{k+2}\right\}.$$

The result then follows.

Case 4 Denote $s_i^2 = p_i(1-p_i), i = 1, \dots, d$. Since $(r_1 + \cdots + r_k)/n + r_{k+1}/(2n)$ converges to a normal distribution with mean $p_1 + \cdots + p_k + p_{k+1}/2 = \pi_{k+1} = p$, variance $O(1/n)$, hence for n sufficiently large, we have either

$$p = \lambda_{n1}\left(\frac{r_1 + \cdots + r_{k-1}}{n} + \frac{r_k}{2n}\right) + (1 - \lambda_{n1})\left(\frac{r_1 + \cdots + r_k}{n} + \frac{r_{k+1}}{2n}\right)$$

if $(r_1 + \cdots + r_k)/n + r_{k+1}/(2n) \geq p$ or

$$p = \lambda_{n2}\left(\frac{r_1 + \cdots + r_{k+1}}{n} + \frac{r_{k+2}}{2n}\right) + (1 - \lambda_{n2})\left(\frac{r_1 + \cdots + r_k}{n} + \frac{r_{k+1}}{2n}\right),$$

if $(r_1 + \cdots + r_k)/n + r_{k+1}/(2n) < p$. Here, $0 \leq \lambda_{n1}, \lambda_{n2} < 1$. We consolidate the two cases into the form

$$\begin{aligned} p = \frac{\sum_{i=1}^k r_i}{n} + \frac{r_{k+1}}{2n} - I\left(\frac{\sum_{i=1}^k r_i}{n} + \frac{r_{k+1}}{2n} \geq p\right)\frac{\lambda_{n1}}{2n}(r_k + r_{k+1}) \\ + I\left(\frac{\sum_{i=1}^k r_i}{n} + \frac{r_{k+1}}{2n} < p\right)\frac{\lambda_{n2}}{2n}(r_{k+1} + r_{k+2}), \end{aligned}$$

therefore, denoting $U = \sqrt{n}\{p - \sum_{i=1}^k r_i/n - r_{k+1}/(2n)\}$, we have

$$\begin{aligned} \hat{\xi} &= I(U \leq 0)\{\lambda_{n1}v_k + (1-\lambda_{n1})v_{k+1}\} + I(U > 0)\{\lambda_{n2}v_{k+2} + (1-\lambda_{n2})v_{k+1}\} \\ &= v_{k+1} + \frac{U}{\sqrt{n}}\left\{\frac{v_{k+1} - v_k}{(r_k + r_{k+1})/(2n)} I(U \leq 0) + \frac{v_{k+2} - v_{k+1}}{(r_{k+1} + r_{k+2})/(2n)} I(U > 0)\right\}. \end{aligned}$$

Thus, for any $a \leq 0$, we have

$$\Pr\{\sqrt{n}(\hat{\xi} - v_{k+1}) \leq a\} = \Pr\left\{U\frac{v_{k+1} - v_k}{(r_k + r_{k+1})/(2n)} \leq a\right\}.$$

Similar to case 3, we have

$$U = -\sum_{i=1}^{k} s_i Z_i - s_{k+1} Z_{k+1}/2,$$
$$(r_k + r_{k+1})/(2n) = (p_k + p_{k+1})/2 + n^{-1/2}(s_k Z_k + s_{k+1} Z_{k+1})/2,$$
$$(r_{k+2} + r_{k+1})/(2n) = (p_{k+2} + p_{k+1})/2 + n^{-1/2}(s_{k+2} Z_{k+2} + s_{k+1} Z_{k+1})/2,$$

where Z_i's are defined as in case 3. Thus,

$$\Pr\{\sqrt{n}(\hat{\xi} - v_{k+1}) \le a\} \to \Pr\left\{\frac{-\sum_{i=1}^{k} 2s_i Z_i - s_{k+1} Z_{k+1}}{p_k + p_{k+1}} \le \frac{a}{v_{k+1} - v_k}\right\}.$$

Obviously, $(-\sum_{i=1}^{k} 2s_i Z_i - s_{k+1} Z_{k+1})/(p_k + p_{k+1})$ converges to a normal random variable with mean zero and variance σ_1^2, where

$$\begin{aligned}\sigma_1^2 &= (p_k + p_{k+1})^{-2}\left(\sum_{i=1}^{k} 4s_i^2 + s_{k+1}^2 - 8\sum_{i,j=1,i\neq j}^{k} p_i p_j - 4p_{k+1}\sum_{i=1}^{k} p_i\right)\\ &= (p_k + p_{k+1})^{-2}\left\{\sum_{i=1}^{k} 4p_i + p_{k+1} - \left(\sum_{i=1}^{k} 2p_i + p_{k+1}\right)^2\right\}\\ &= (p_k + p_{k+1})^{-2}\,(4p(1-p) - p_{k+1}).\end{aligned}$$

Hence we obtain

$$\Pr\{\sqrt{n}(\hat{\xi} - v_{k+1}) \le a\} \to \Phi\left\{\frac{a}{\sigma_1(v_{k+1} - v_k)}\right\}$$

when $n \to \infty$, the result then follows for $\hat{\xi} > v_{k+1}$.

For any $a > 0$, we have

$$\begin{aligned}\Pr\{\sqrt{n}(\hat{\xi} - v_{k+1}) \le a\} &= \Pr\left\{U \le 0 \text{ or } 0 < U\frac{v_{k+2} - v_{k+1}}{(r_{k+2} + r_{k+1})/(2n)} \le a\right\}\\ &= \Pr\left\{U\frac{v_{k+2} - v_{k+1}}{(r_{k+2} + r_{k+1})/(2n)} \le a\right\}\\ &\to \Pr\left\{\frac{-\sum_{i=1}^{k} 2s_i Z_i - s_{k+1} Z_{k+1}}{p_{k+2} + p_{k+1}} \le \frac{a}{v_{k+2} - v_{k+1}}\right\}.\end{aligned}$$

Similarly $(-\sum_{i=1}^{k} 2s_i Z_i - s_{k+1} Z_{k+1})/(p_{k+2} + p_{k+1})$ converges to a normal random variable with mean zero and variance σ_2^2, where

$$\sigma_2^2 = (p_{k+1} + p_{k+2})^{-2}\,(4p(1-p) - p_{k+1}).$$

We thus obtain

$$\Pr\{\sqrt{n}(\hat{\xi} - v_{k+1}) \leq a\} \to \Phi\left\{\frac{a}{\sigma_2(v_{k+2} - v_{k+1})}\right\}$$

when $n \to \infty$, the result then follows for $\hat{\xi} > v_{k+1}$.

Combining the above results, we obtain the proof. □

Remark 2 From the main statement and proof of Theorem 2, it is clear that d does not have to be a finite number. When $d = \infty$, case 2 will be eliminated, and all other cases remain the same.

Remark 3 The result in case 3 of Theorem 2 is in accordance with the result in case 1 of Theorem 1. It is not difficult to convince ourselves that $\lambda v_{k+1} + (1 - \lambda)v_{k+2}$ is a reasonable candidate for the "true" quantile ξ. The term $-\{1 - (\lambda - 1)^2\}p_{k+1}/4 - (1 - \lambda^2)p_{k+2}/4$ that we subtract from $p(1 - p)$ can be viewed as an efficiency gain from the fact that the mass is more concentrated in the discrete case than in the continuous case. This term vanishes in the continuous case because there $p_{k+1} = p_{k+2} = 0$.

Remark 4 Similarly, the result in case 4 of Theorem 2 is in accordance with the result in case 2 of Theorem 1.

Remark 5 The result of case 2 in Theorem 1 can be equivalently expressed as

$$\sqrt{n}(\hat{\xi} - \xi) \to N_-\left\{0, \frac{p(1-p)}{f_-^2(\xi)}\right\} + N_+\left\{0, \frac{p(1-p)}{f_+^2(\xi)}\right\}$$

and the result in case 4 of Theorem 2 can be equivalently expressed as

$$\sqrt{n}(\hat{\xi} - v_{k+1}) \to N_-\left\{0, \frac{p(1-p) - p_{k+1}/4}{f_-^2(v_{k+1})}\right\} + N_+\left\{0, \frac{p(1-p) - p_{k+1}/4}{f_+^2(v_{k+1})}\right\}$$

in distribution as $n \to \infty$. Here, $N_-(0, \sigma_-^2) + N_+(0, \sigma_+^2)$ represents a distribution whose density function is that of a normal with variance σ_-^2 on the negative real line and that of a normal with variance σ_+^2 on the positive real line. Such a distribution is termed a joined half-Gaussian or two-piece normal distribution (see Gibbons and Mylroie 1973; John 1982). Case 2 provides one of the rare "natural" situations where such a distribution arises in a limiting process.

Remark 6 Note that the results in case 2 of Theorem 1 and in case 4 of Theorem 2 indicate that the asymptotic convergence of sample quantiles based on mid-distribution functions is not the typical root-n rate, but of a slightly slower order $o_p(\sqrt{n})$. This is because the combination of the two normals typically does not yield a zero mean. Instead, it has mean $\frac{\sigma_+ - \sigma_-}{\sqrt{2\pi}}$, where σ_+, σ_- are the standard deviations of the underlying normal distributions for the positive and negative region, respectively.

Remark 7 Many distributions consist of both discrete parts and continuous parts. Since the asymptotic property of a certain sample quantile only concerns the local properties of the distribution, one can typically adapt the results from Theorems 1 and 2 to obtain the corresponding results in different regions of the distribution. A special case, for example, is when the probability density function has support consisting of several disjoint intervals and some discrete points. Thus, the corresponding cumulative distribution function F has differentiability in certain regions and also discontinuity points. While both are easily treated by either Theorem 1 or 2, there is also a third type of points that involve differentiability from one side and discontinuity from the other side. These are therefore not readily solvable using any of the theorems. However, it can be easily seen that the asymptotic result in that situation is a combination of case 2 of Theorem 1 and case 4 of Theorem 2.

5 Exact results for the binomial and Poisson distributions

In practice, binomial and Poisson types of data are frequently encountered. We derive the exact distribution of the sample quantiles defined through mid-distribution functions in those two cases for illustration. The asymptotic distribution is covered by Theorem 2.

Assume $X_1, \ldots, X_n$ are observations from a binomial distribution Binomial(m, q_0) with values $0, 1, \ldots, m$. Suppose the count of 0 among the n observations is r_0, and similarly, there are r_1 1's, ..., r_m m's. Note that $n = \sum_{i=0}^{m} r_i$. According to the calculations in Sect. 2, the sample quantiles for a specific value $0 \le p \le 1$ are

$$\tilde{Q}(p) = \begin{cases} 0 & \text{if } p \le \frac{r_0}{2n}, \\ m & \text{if } p \ge 1 - \frac{r_m}{2n}, \\ k+1-\lambda & \text{if } p = \lambda\left(\dfrac{r_0+\cdots+r_{k-1}}{n} + \dfrac{r_k}{2n}\right) + (1-\lambda) \\ & \quad \times \left(\dfrac{r_0+\cdots+r_k}{n} + \dfrac{r_{k+1}}{2n}\right), k = 0, \ldots, m-1. \end{cases} \tag{4}$$

Therefore, denoting $p_k = \binom{m}{k} q_0^k (1-q_0)^{m-k}$ for $k = 0, \ldots, m$, the exact distribution of $\tilde{Q}(p)$ is

$$P\left\{\tilde{Q}(p) = 0\right\} = \sum_{k \ge 2np} \binom{n}{k} p_0^k (1-p_0)^{n-k}, \tag{5}$$

$$P\left\{\tilde{Q}(p) = m\right\} = \sum_{k \ge 2n-2np} \binom{n}{k} p_m^k (1-p_m)^{n-k}, \tag{6}$$

$$P\left\{\tilde{Q}(p) = k + 1 - \frac{2\sum_{i=0}^{k} c_i + c_{k+1} - 2np}{c_k + c_{k+1}}\right\}$$

$$= \sum_{c_0,\ldots,c_{k+1}} \binom{n}{c_0,\ldots,c_{k+1}} p_0^{c_0} \cdots p_{k+1}^{c_{k+1}} \left(1 - \sum_{i=0}^{k+1} p_i\right)^{n-\sum_{i=0}^{k+1} c_i}$$

for integers $k, c_0, \ldots, c_{k+1}$ such that

$$0 \le 2\sum_{i=0}^{k} c_i + c_{k+1} - 2np \le c_k + c_{k+1}. \tag{7}$$

Similar results hold for the Poisson distribution with parameter μ. Assume we have n observations among which there are r_0 0's, r_1 1's, ..., r_m m's, where m can be an arbitrary large integer. Note that here also $n = \sum_{i=0}^{m} r_i$. We can easily obtain that for any $0 \le p \le 1$, $\tilde{Q}(p)$ is given by (4) and its exact distribution is (5)–(7), where now $p_k = \mu^k e^{-\mu}/k!$ for $k = 0, 1, \ldots, \infty$.

6 Simulation studies

We perform a small simulation study to verify the asymptotic results established in Theorems 1 and 2. For the results regarding the continuous situation, we generate data from a combination of three gamma distributions. Specifically, observations in the interval (0, 1.4452) are generated from a gamma distribution with parameters $\alpha = 3.6$ and $\beta = 0.5970$; observations in the interval (1.4452, 4.8784) are generated from a gamma distribution with parameters $\alpha = 1$ and $\beta = 4.0519$; and finally observations in the interval $(4.8784, \infty)$ are generated from a gamma distribution with parameters $\alpha = 2$ and $\beta = 2$. We choose such a data generation pattern because the resulting cumulative distribution function has both differentiable points and one-sided differentiable points. Specifically, for the 0.3th and 0.7th quantile, the cumulative distribution function is only differentiable from one side, while for all other quantiles, it is differentiable. We calculate the sample quantiles based on mid-distribution functions for probabilities $p = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$. A sample size of $n = 100$ is used and we repeat the simulation 1,000 times. The sample mean and sample variance of the sample quantiles, the theoretical quantiles and the asymptotic variance calculated according to Theorem 1 are presented in Table 1. It is worth pointing out that at the $p = 0.3$ and $p = 0.7$ quantiles, the asymptotic mean is not simply $F^{-1}(p)$, but has a slightly biased form $F^{-1}(p) + (\sigma_+ - \sigma_-)/\sqrt{2\pi n}$, where σ_+ and σ_- are the asymptotic standard deviations from the right and left sides, respectively. As a consequence, the asymptotic variance is calculated as $(\sigma_+^2 + \sigma_-^2)/2n - \{(\sigma_+ - \sigma_-)/\sqrt{2\pi n}\}^2$. We also calculated the 95% confidence interval and reported the coverage. Here, for $p = 0.3$ and $p = 0.7$, the 95% confidence interval is an asymmetric one, with left and right end points equaling the 0.025 quantile of the centered normal distribution with variance σ_-^2 and 0.975 quantile of the centered normal distribution with variance σ_+^2, respectively. As is clear from the numerical results, the asymptotic properties are verified quite well even for a small sample size of $n = 100$.

The second simulation is performed to verify the results in Theorem 2. We generate data from a discrete distribution with point mass at −2.3, −0.5, 0, 1, 2.5 and equal

Table 1 Simulation results for the continuous distribution. Data generated from a combination of three gamma distributions

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
ξ	0.8843	1.1866	1.5052	2.0698	2.8086	3.7128	4.8034	5.9886	7.7794
ave($\hat{\xi}$)	0.8873	1.1852	1.5062	2.0656	2.8152	3.7233	4.8088	5.9888	7.7629
$n$var($\hat{\xi}$)	1.0970	1.2196	4.2253	10.3789	16.3905	25.0669	27.0467	30.3886	56.0074
$n\hat{\sigma}^2$	1.1116	1.1791	3.8185	10.9455	16.4183	24.6274	27.8683	28.4715	56.8897
95%	0.9600	0.9550	0.9470	0.9600	0.9530	0.9480	0.9450	0.9400	0.9490

In $(0, 1.4452)$, the gamma distribution has parameters $\alpha = 3.6$ and $\beta = 0.5970$; in $(1.4452, 4.8784)$, the gamma distribution has parameters $\alpha = 1$ and $\beta = 4.0519$; in $(4.8784, \infty)$, the gamma distribution has parameters $\alpha = 2$ and $\beta = 2$. $\hat{\sigma}^2$ represents the estimated variance of $\hat{\xi}$, '95%' represents the 95% confidence interval coverage probability. Results are based on 1,000 simulations, with sample size 100

Table 2 Simulation results for the discrete distribution

p	1/15	1/10	3/10	3/5	3/4	9/10	15/16
ξ	−2.3000	−2.2679	−0.5464	0.5000	1.3750	2.4732	2.5000
ave($\hat{\xi}$)	−2.3000	−2.2661	−0.5402	0.4997	1.3709	2.4735	2.5000
$n$var($\hat{\xi}$)	0	1.1488	4.9072	3.8968	6.4365	0.8163	0
$n\hat{\sigma}^2$	0	1.1043	5.9041	4.1250	6.6797	0.7669	0
95%	N/A	0.9460	0.9690	0.9540	0.9590	0.9430	N/A

Data generated from a discrete distribution with equal weights and point mass at $-2.3, -0.5, 0, 1, 2.5$. ξ represents the theoretical convergence center of the estimated $\hat{\xi}$'s, $\hat{\sigma}^2$ represents the estimated variance of $\hat{\xi}$, '95%' represents the 95% confidence interval coverage probability. Results are based on 1,000 simulations, with sample size 500

weights 1/5. We calculate the sample quantiles based on mid-distribution functions for various probabilities, including $p = 1/15, 1/10, 3/10, 3/5, 3/4, 9/10, 15/16$. We present the sample mean and sample variance of the estimates, as well as the theoretical mean and variance from Theorem 2. Note that for cases 1 and 2, which correspond to $p = 1/15$ and $p = 15/16$, the theoretical variance is simply zero. For case 3, which corresponds to $p = 3/5$ and $p = 3/4$, the asymptotic variance and confidence interval can be calculated using the asymptotic normality result. For case 4, which corresponds to $p = 1/10$, $p = 3/10$ and $p = 9/10$, the theoretical mean is calculated as $\xi = v_{k+1} + (\sigma_+ - \sigma_-)/\sqrt{2\pi n}$, and the theoretical variance is calculated as $(\sigma_-^2 + \sigma_+^2)/(2n) - \{(\sigma_+ - \sigma_-)/\sqrt{2\pi n}\}^2$. Here, σ_-^2 and σ_+^2 are the asymptotic variances of the negative and positive sides, respectively. In addition, the 95% confidence interval for case 4 is also quite specific. For $p = 3/10$, we calculated the asymmetric confidence interval, with the left end point corresponding to the 0.025 quantile of the $N(0, \sigma_-^2)$ distribution and the right end point corresponding to the 0.975 quantile of the $N(0, \sigma_+^2)$ distribution. For $p = 1/10$ and $p = 9/10$, because one side of the final distribution deteriorates to a point mass at 0 with weight 0.5, we calculated the 95% confidence interval using 0 as one end point. The other end point of the confidence interval is taken to be the 0.95 quantile of the $N(0, \sigma_+^2)$ distribution for $p = 1/10$ and the 0.05 quantile of the $N(0, \sigma_-^2)$ distribution for $p = 9/10$. The results based on sample size $n = 500$ and 1,000 simulations are reported in Table 2. The

Table 3 Simulation results for the binomial distribution

p	1/15	1/10	3/10	3/5	3/4	9/10	15/16
ξ	0	0	0.4672	1.2662	1.6665	2	2
ave($\hat{\xi}$)	0	0	0.4667	1.2667	1.6667	2.0000	2.0000
$n$var($\hat{\xi}$)	0	0.0001	0.4883	0.6568	0.4352	0.0002	0
$n\hat{\sigma}^2$	0	0	0.5096	0.6756	0.4444	0	0
95%	N/A	N/A	0.9590	0.9590	0.9550	N/A	N/A

Data generated from a binomial distribution with $m = 2$ and $q_0 = 0.5$. ξ represents the theoretical convergence center of the estimated $\hat{\xi}$'s, $\hat{\sigma}^2$ represents the estimated variance of $\hat{\xi}$, '95%' represents the 95% confidence interval coverage probability. Results are based on 1,000 simulations, with sample size 500

Table 4 Simulation results for the Poisson distribution

p	1/15	1/10	3/10	3/5	3/4	9/10	15/16
ξ	0	0.0014	0.6776	1.7164	2.3528	3.3225	3.7610
ave($\hat{\xi}$)	0	0	0.6756	1.7154	2.3541	3.3291	3.7637
$n$var($\hat{\xi}$)	0	0.0154	1.2453	1.5294	3.0541	7.3237	4.5283
$n\hat{\sigma}^2$	0	0	1.3473	1.6496	3.2255	7.4607	4.6210
95%	N/A	N/A	0.9620	0.9580	0.9530	0.9430	0.9470

Data generated from a Poisson distribution with $\mu = 1.5$. ξ represents the theoretical convergence center of the estimated $\hat{\xi}$'s, $\hat{\sigma}^2$ represents the estimated variance of $\hat{\xi}$, '95%' represents the 95% confidence interval coverage probability. Results are based on 1,000 simulations, with sample size 500

numerical results show that for sufficiently large sample size, the asymptotic results are relevant.

To further demonstrate the discrete case, we conducted simulations under binomial and Poisson data with sample size $n = 500$. For the binomial data, the true distribution that generates the data is Binomial($m = 2, q_0 = 0.5$) and therefore the three supporting values are 0, 1 and 2. For the Poisson data, the true parameter is $\mu = 1.5$. The remaining set up is the same as in the second simulation. We summarize the outcomes in Tables 3 and 4 which again show that for sufficiently large sample size, the asymptotic results are relevant.

Acknowledgments The authors would like to thank the Editor, the Associate Editor, and two referees for comments that improved this manuscript. This research was partially supported by NSF grants DMS-0504896 and CMG ATM-0620624.

References

Berry, G., Armitage, P. (1995). Mid-*P* confidence intervals: a brief review. *The Statistician*, *44*, 417–423.

David, H. A., Nagaraja, H. N. (2003). *Order statistics* (3rd ed.). Hoboken: Wiley.

Galton, F. (1889). *Natural inheritance*. London: Macmillan.

Genest, C., Nešlehová, J. (2007). A primer on copulas for count data. *The Astin Bulletin*, *37*, 475–515.

Genton, M. G., Ma, Y., Parzen, E. (2006). Discussion of "Sur une limitation très générale de la dispersion de la médiane" by M. Fréchet, 1940. *Journal de la Société Française de Statistique*, *147*, 51–60.

Gibbons, J. F., Mylroie, S. (1973). Estimation of impurity profiles in ion-implanted amorphous targets using joined half-Gaussian distributions. *Applied Physics Letters*, *22*, 568–569.

González-Barrios, J., Rueda, R. (2001). On convergence theorems for quantiles. *Communications in Statistics Theory and Methods*, *30*, 943–955.

Hald, A. (1998). *A history of mathematical statistics from 1750 to 1930*. New York: Wiley.

John, S. (1982). The three-parameter two-piece Normal family of distributions and its fitting. *Communications in Statistics Theory and Methods*, *11*, 879–885.

Lancaster, H. O. (1961). Significance tests in discrete distributions. *Journal of the American Statistical Association*, *56*, 223–234.

Larocque, D., Randles, R. H. (2008). Confidence intervals for a discrete population median. *The American Statistician*, *62*, 32–39.

Machado, J., Santos Silva, J. (2005). Quantiles for counts. *Journal of the American Statistical Association*, *100*, 1226–1237.

Parzen, E. (1954). *On uniform convergence of families of sequences of random variables*. California Publications in Statistics, Berkeley: University of California Press.

Parzen, E. (1997). Concrete statistics. In: Ghosh, S., Schucany, W. R., Smith, W. B. (Eds.), *Statistics in quality* (pp. 309–332). New York: Marcel Dekker.

Parzen, E. (2004). Quantile probability and statistical data modeling. *Statistical Science*, *19*, 652–662.

Smirnov, N. V. (1952). Limit distributions for the terms of a variational series. *American Mathematical Society*, Ser. 1, Translation Number *67*, 1–64.

Walker, A. M. (1968). A note on the asymptotic distribution of sample quantiles. *Journal of the Royal Statistical Society Series B*, *30*, 570–575.