Journal of Econometrics 185 (2015) 20–32

Contents lists available at ScienceDirect

Journal of Econometrics

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

Improved quantile inference via fixed-smoothing asymptotics and Edgeworth expansion

David M. Kaplan

Department of Economics, University of Missouri, 909 University Ave, 118 Professional Bldg, Columbia, MO 65211-6040, United States

ARTICLE INFO

Article history:
Received 26 July 2012
Received in revised form
14 June 2014
Accepted 10 August 2014
Available online 30 October 2014

JEL classification:
C21

Keywords:
Fixed-smoothing
High-order accuracy
Hypothesis testing
Testing-optimal smoothing parameter

ABSTRACT

To estimate a sample quantile's variance, the quantile spacing method involves smoothing parameter m. When $m, n \to \infty$, the corresponding Studentized test statistic is asymptotically $N(0, 1)$. Holding m fixed instead, the asymptotic distribution contains the Edgeworth expansion term capturing the variance of the quantile spacing. Consequently, the fixed-m distribution is more accurate than the standard normal under both asymptotic frameworks. A testing-optimal m is proposed to maximize power subject to size control. In simulations, the new method controls size better than similar methods while maintaining good power. Throughout are results for two-sample quantile treatment effect inference. Code is available online.

1. Introduction

This paper considers inference on population quantiles, specifically via the Studentized test statistic from Siddiqui (1960) and Bloch and Gastwirth (1968) (jointly SBG hereafter), as well as inference on quantile treatment effects in a two-sample (treatment/control) setup. Median inference is a special case. Nonparametric inference on conditional quantiles is an immediate extension if the conditioning variables are all discrete. In addition to common variables like race and sex, variables like years of education may be treated as discrete as long as enough observations exist for each value of interest. Continuous conditioning variables are accommodated by the approach in Goldman and Kaplan (2014a), whose code includes an implementation of this paper's method. Like a t-statistic for the mean, the SBG statistic is normalized using a consistent estimate of the variance of the quantile estimator, and it is asymptotically standard normal.

Quantile treatment effects can enrich the usual average treatment effect analysis in economic experiments, such as those in Björkman and Svensson (2009), Charness and Gneezy (2009), and Gneezy and List (2006). Quantile treatment effects have also been discussed recently in Bitler et al. (2006) for welfare programs, in Djebbari and Smith (2008) for the PROGRESA conditional cash transfer program in Mexico, and in Jackson and Page (2013) for heterogeneous effects of class size in the Tennessee STAR program. Quantiles have proved of interest across many other fields within economics, such as health (Abrevaya, 2001), labor (Angrist et al., 2006; Buchinsky, 1994), auctions (Guerre and Sabbah, 2012), and land valuation (Koenker and Mizera, 2004). In finance, value-at-risk (VaR) is defined as a quantile; this paper's method can construct a confidence interval for the VaR, using a "historical simulation" approach or that in Section 3 of Cabedo and Moya (2003). In insurance, the Scenario Upper Loss (SUL) is an upper quantile of the property loss distribution given an upper quantile magnitude earthquake; the less consistently defined probable maximum loss (PML) is similar and also widely used.

I develop new asymptotic theory that is more accurate than the standard normal reference distribution traditionally used with the SBG statistic. With more accuracy comes improved inference properties, i.e. controlling size where previous methods were size-distorted without sacrificing much power. A plug-in procedure to choose the testing-optimal smoothing parameter m translates the theory into practice. Inverting the level-α tests proposed here yields level-α confidence intervals. I use hypothesis testing language throughout, but size distortion is analogous to coverage probability error, and higher power corresponds to shorter interval lengths. It is also straightforward to solve numerically for p-values using the higher-order critical value approximation since it is a simple arithmetic correction of standard normal critical values.

The two key results here are a nonstandard "fixed-m" asymptotic distribution (where m is a smoothing parameter used for Studentization) and a high-order Edgeworth expansion. For a scalar

E-mail address: kaplandm@missouri.edu.

http://dx.doi.org/10.1016/j.jeconom.2014.08.011
0304-4076/

location model, Siddiqui (1960) gives the fixed-m result, and Hall and Sheather (1988, hereafter cited as "HS88") give a special case of the Edgeworth expansion in Theorem 2 below. The Edgeworth expansion is more accurate than a standard normal since it contains high-order terms that are otherwise ignored.

In the standard "large-m" asymptotics, both $m \to \infty$ and $n \to \infty$ (and $m/n \to 0$). In contrast, fixed-m asymptotics only approximates $n \to \infty$ while fixing m at its actual finite-sample value. It turns out that the fixed-m asymptotics includes the high-order Edgeworth term capturing the variance of the quantile spacing in the Studentized quantile's denominator. (Fixed-m is an instance of "fixed-smoothing" asymptotics in that this variance does not go to zero in the limit as in the "increasing-smoothing" large-m asymptotics.) Thus, from the theoretical view, the fixed-m distribution is more accurate than the standard normal irrespective of the asymptotic framework: it is high-order accurate under large-m, while the standard normal approximation is not even first-order accurate under fixed-m.

The Edgeworth and fixed-m distributions also capture the effect of the choice of m, whereas the standard normal approximation does not. I construct a test dependent on m using the fixed-m critical values and evaluate the type I and II error rates of the test using the more accurate Edgeworth expansion. Then I optimally select m to minimize type II error subject to control of type I error.

This work builds partly on Goh (2004), who suggests fixed-m asymptotics for Studentized quantile inference. Goh (2004) uses simulated critical values and prior choices of m. Here, I provide a simple, accurate fixed-m critical value approximation and corresponding new optimal choice of m. I also show the improved theoretical accuracy of the fixed-m distribution, complementing the simulations in Goh (2004).

In the time series context of heteroskedasticity–autocorrelation robust inference, the key ideas of "testing-optimal" smoothing parameter choice, fixed-smoothing (or "fixed-b") asymptotics, higher-order asymptotics, and corrected critical values based on a common distribution appear in a sequence of papers by Sun et al. (2008), Sun (2013, 2011, 2014), and Sun and Kaplan (2011). In particular, analogous to the fixed-m results below, Sun et al. (2008) show that their testing-optimal bandwidth is of a different order than the MSE-optimal bandwidth, and Sun (2011) generalizes the result from Sun et al. (2008) that shows fixed-smoothing asymptotics to be a higher-order refinement of increasing-smoothing asymptotics.

Whereas in HS88 the choice of m is critical to controlling size (or not), the fixed-m critical values provide size control robust to incorrectly chosen m. In simulations, the new method has correct size even where the HS88 method is size-distorted. Power is still good because m is explicitly chosen to maximize it, using the Edgeworth expansion. HS88 do not provide a separate result for the two-sample (quantile treatment effect) case.

Monte Carlo simulations show that the new method controls size better than HS88 and various bootstrap methods, while maintaining competitive power. Following a fractional order statistic approach, the one-sample method of Hutson (1999), shown to have $O(n^{-1})$ coverage probability error by Goldman and Kaplan (2014a), and the related two-sample method of Goldman and Kaplan (2014b) usually have better properties but are not always computable. Thus, they complement the new method herein. Additionally, they produce equal-tailed confidence intervals, while the new method's intervals are symmetric; preference may depend on the application. The new method also has a computational advantage over both methods, especially the two-sample method.

The basic setup is in Section 2. Sections 3 and 4 concern fixed-m and Edgeworth results, respectively. A consequently testing-optimal choice of m is proposed in Section 5. Section 6 contains simulation results. Two-sample results are provided in parallel. More details are in the working paper version, and full technical proofs and calculations are in the supplemental appendices (see Appendix E). These and computer code are available on the author's website.

2. Quantile estimation and hypothesis testing

Consider an iid sample of continuous random variable X, whose p-quantile is ξ_p. The estimator $\hat{\xi}_p$ used in SBG is an order statistic. The rth order statistic for a sample of n values is the rth smallest value in the sample, $X_{n,r}$, such that $X_{n,1} < X_{n,2} < \cdots < X_{n,n}$. The SBG estimator is

$$\hat{\xi}_p = X_{n,r}, \quad r = \lfloor np \rfloor + 1,$$

where $\lfloor np \rfloor$ is the floor function. Writing the cumulative distribution function (CDF) of X as $F(x)$ and the probability density function (PDF) as $f(x) \equiv F'(x)$, I make the usual assumption that $f(x)$ is positive and continuous in a neighborhood of the point ξ_p. Consequently, ξ_p is the unique p-quantile such that $F(\xi_p) = p$.

The standard asymptotic result (Mosteller, 1946; Siddiqui, 1960) for a central[1] quantile estimator is

$$\sqrt{n}(X_{n,r} - \xi_p) \xrightarrow{d} N\left(0,\ p(1-p)\left[f(\xi_p)\right]^{-2}\right). \tag{1}$$

A consistent estimator of $1/f(\xi_p)$ that is asymptotically independent of $X_{n,r}$ leads to the Studentized sample quantile, which has the pivotal asymptotic distribution

$$\frac{\sqrt{n}(X_{n,r} - \xi_p)}{\sqrt{p(1-p)}\left[\widehat{1/f(\xi_p)}\right]} \xrightarrow{d} N(0, 1). \tag{2}$$

Siddiqui (1960) and Bloch and Gastwirth (1968) propose and show consistency of

$$\widehat{1/f(\xi_p)} = S_{m,n} \equiv \frac{n}{2m}(X_{n,r+m} - X_{n,r-m}) \tag{3}$$

when $m \to \infty$ and $m/n \to 0$ as $n \to \infty$.

For two-sided inference on ξ_p, I consider the parameterization $\xi_p = \beta - \gamma/\sqrt{n}$. The null and alternative hypotheses are $H_0 : \xi_p = \beta$ and $H_1 : \xi_p \neq \beta$, respectively. When $\gamma = 0$, the null is true. The test statistic examined in this paper is

$$T_{m,n} \equiv \frac{\sqrt{n}(X_{n,r} - \beta)}{S_{m,n}\sqrt{p(1-p)}} \tag{4}$$

and will be called the SBG test statistic due to its use of (3). From (2), $T_{m,n}$ is asymptotically standard normal when $\gamma = 0$. The corresponding hypothesis test compares $T_{m,n}$ to critical values from a standard normal distribution.

For the two-sample case, assume that there are independent samples of X and Y, with n_x and n_y observations, respectively. For simplicity, let $n = n_x = n_y$. For instance, if $2n$ individuals are separated into balanced treatment and control groups, one might want to test if the treatment effect at quantile p has significance level α. The marginal PDFs are $f_X(\cdot)$ and $f_Y(\cdot)$. Under the null hypothesis $H_0 : \xi_{px} = \xi_{py} = \xi_p$,

$$\sqrt{n}(X_{n,r} - \xi_p) - \sqrt{n}(Y_{n,r} - \xi_p) \\ \xrightarrow{d} N\left(0, p(1-p)([f_X(\xi_p)]^{-2} + [f_Y(\xi_p)]^{-2})\right),$$

[1] "Central" means that in the limit, $r/n \to p \in (0, 1)$ as $n \to \infty$; i.e., $r \to \infty$ and is some fraction of the sample size n. In contrast, "intermediate" would take $r \to \infty$ but $r/n \to 0$ (or $r/n \to 1$, $n - r \to \infty$); "extreme" would fix $r < \infty$ or $n - r < \infty$.