

Testing Equality of Nonparametric Quantile Regression Functions

Chinthaka Kuruwita¹, Colin Gallagher² & K. B. Kulasekera³

¹ Department of Mathematics, Hamilton College, Clinton, USA

² Department of Mathematical Sciences, Clemson University, Clemson, USA

³ Department of Bioinformatics & Biostatistics, University of Louisville, USA

Correspondence: Chinthaka Kuruwita, Department of Mathematics, Hamilton College, Clinton, NY 13323, USA.
Tel: 1-315-859-4365. E-mail: ckuruwit@hamilton.edu

Received: November 14, 2013 Accepted: December 15, 2013 Online Published: January 6, 2014

doi:10.5539/ijsp.v3n1p55 URL: <http://dx.doi.org/10.5539/ijsp.v3n1p55>

Abstract

This article proposes a new approach for testing the equality of nonparametric quantile regression functions based on marked empirical processes. We develop test statistics that possess better Type I and power properties in comparison to all available procedures in the literature. Simulation results also indicate that our tests have superior local power properties over existing tests. A data analysis is given which highlights the usefulness of the proposed methodology.

Keywords: robust regression, empirical process, local linear estimator

1. Introduction

Comparing groups based on independent samples has been a fundamental problem in statistics. In the context of regression analysis, this is typically done by comparing the conditional means of the groups. In parametric regression, comparison of model parameters of a predefined parametric model provides information about differences among the groups. When a parametric specification is not appropriate, one compares nonparametric mean regression functions to detect any differences between groups (Hall & Hart, 1990; King, Hart, & Wehrly, 1991; Kulasekera, 1995; Neumeyer & Dette, 2003). The above mentioned procedures rely on the assumption that the errors of the assumed model have finite variance. In practice this assumption may not hold, especially with heavy tailed error distributions. Furthermore, these procedures are highly sensitive to extreme observations which can lead to spurious results. To overcome these difficulties in mean regression procedures, Koenker and Bassett (1978) in their pioneering work proposed the quantile regression framework where, one regresses the conditional quantiles of a response on covariates rather than regressing the conditional mean. In this article we develop a flexible and robust testing procedure to compare groups within the quantile regression framework. Our method uses nonparametric quantile regression functions to construct test statistics that can detect differences at targeted quantiles in two or more conditional distributions.

Suppose we observe data from k independent groups in the form $(X_{ij}, Y_{ij}), i = 1, \dots, k, j = 1, \dots, n_i$, where X is a continuous covariate supported on $[0, 1]$ and Y is a continuous response. For $\tau \in (0, 1)$, let $g_\tau(X)$ denote the τ^{th} conditional quantile function of Y given X (i.e., $P[Y \leq g_\tau(X)|X] = \tau$). We model $g_\tau(\cdot)$ nonparametrically with the following nonparametric quantile regression model:

$$Y_{ij} = g_{\tau,i}(X_{ij}) + \epsilon_{ij}, \quad i = 1, \dots, k, \quad j = 1, \dots, n_i \quad (1)$$

where $g_{\tau,i}$ is the τ^{th} conditional quantile function of Y given X for the i^{th} group and ϵ_{ij} is a sequence of independent random variables assumed to be identically distributed within each group. We further assume that the τ^{th} conditional quantile of ϵ_{ij} given X_{ij} is zero. We are interested in testing whether the conditional quantile curves are the same across the k groups or not for a fixed $\tau \in (0, 1)$. That is, testing

$$H_0 : g_{\tau,1}, \dots, g_{\tau,k} = g_\tau \quad \text{vs} \quad H_1 : g_{\tau,i} \neq g_{\tau,i'} \quad \text{for some } i \neq i'$$

over the support of the covariate X .

Sun (2006) and Dette, Wagner, and Volgushev (2011) have compared nonparametric quantile regression functions in the form of the above hypothesis. The former propose a test based on an orthogonal moment condition of residuals which holds under the null hypothesis and the latter uses a test derived from the L^2 distance of the estimated quantile regression curves. Both methods have non-trivial power against local alternatives that converge to the null at a rate $(\sqrt{nh}^{1/4})^{-1}$ with a scalar covariate. Here $n = \sum_{i=1}^k n_i$ is the total sample size and h is the bandwidth (converging to 0) that is used to estimate the nonparametric quantile regression functions assuming the null hypothesis is true. In contrast, we propose a new test procedure that can detect alternatives converging to the null at $n^{-1/2}$ rate. Our method is based on a marked empirical process of residuals. We note that marked empirical processes based on residuals have been used in the context of testing the equality of mean regression functions by Delgado (1993), Kulasekera (1995), and Neumeier and Dette (2003).

The remainder of this article is organized as follows. Section 2 describes the test procedure and presents the bootstrap method that we use to determine the critical values of our test statistics. Section 3 presents the results of the simulation study that investigates the finite sample performance of our method along with a real data example. Concluding remarks are given in Section 4.

2. Construction of Test Statistics

For fixed $\tau \in (0, 1)$ consider the nonparametric quantile regression model in (1). Let $\eta_{ij} = I\{Y_{ij} < g_{\tau,i}(X_{ij})\} - \tau$ and $U_{ij} = I\{Y_{ij} < g_{\tau}(X_{ij})\} - \tau$. Here $I(\cdot)$ denotes the indicator function and g_{τ} is a weighted average of $g_{\tau,i}$ s defined as

$$g_{\tau}(t) = \sum_{i=1}^k \lambda_i(t) g_{\tau,i}(t) \quad (2)$$

with $\lambda_i(t) \in (0, 1)$ and $\sum_{i=1}^k \lambda_i(t) = 1$ (Sun, 2006). Under H_0 , we see that $U_{ij} = \eta_{ij}$ and are independent mean zero random variables. Now consider the process

$$R_i(t) = \frac{1}{n} \sum_{j=1}^{n_i} U_{ij} I(X_{ij} \leq t) \quad (3)$$

for $t \in (0, 1)$. Under H_0 , $R_i(\cdot)$ is a mean zero random process. If the null hypothesis is false, $R_i(\cdot)$ will have a non-zero mean function which can be used as a basis to detect departures from H_0 . To construct test statistics based on process (3), we need an efficient estimator of $R_i(t)$ under H_0 . Estimation of $R_i(t)$ under H_0 solely depends on estimating the quantile regression functions $g_{\tau,i}(\cdot)$. If H_0 is true, all $g_{\tau,i}(\cdot)$'s are the same across the k groups. Therefore, an efficient nonparametric estimator of the common quantile regression function $g_{\tau}(\cdot)$ can be constructed by pooling the data from all groups. For independent data, several nonparametric quantile regression function estimators are available in the literature (Yu & Jones, 1998; Takeuchi, Le, Sears, & Smola, 2006; Dette & Volgushev, 2008; Bondell, Reich, & Wang, 2010). In this paper we use the local linear nonparametric quantile regression function estimator proposed by Yu and Jones (1998). Local linear estimators are popular in practice due to good finite sample performance at boundaries of the support (Fan & Gijbels, 1996). The later estimators are proposed to eliminate the 'quantile crossing' problem in estimating quantile regression functions at multiple quantiles. Since our testing problem targets a fixed $\tau \in (0, 1)$, this problem does not affect our method.

Let $\hat{g}_{\tau}(\cdot)$ be the local linear nonparametric quantile regression function estimator of the common quantile regression function $g_{\tau}(\cdot)$. We define $\hat{g}_{\tau}(x) = \hat{a}$ where \hat{a} and \hat{b} minimize

$$\sum_{i=1}^k \sum_{j=1}^{n_i} \rho_{\tau}(Y_{ij} - a - b(X_{ij} - x)) K\left(\frac{X_{ij} - x}{h}\right). \quad (4)$$

Here $\rho_{\tau}(u) = u\{\tau - I(u \leq 0)\}$ is called the "check function" and h is a smoothing parameter and $K(\cdot)$ is a symmetric density function on $[-1, 1]$. For $i = 1, \dots, k$, define the marked empirical process

$$\hat{R}_i(t) = \frac{1}{n} \sum_{j=1}^{n_i} \hat{U}_{ij} I(X_{ij} \leq t) \quad (5)$$

where $\hat{U}_{ij} = I\{Y_{ij} \leq \hat{g}_{\tau}(X_{ij})\} - \tau$. The process $\hat{R}_i(\cdot)$ in (5) is the building block of our test statistics. Based on these processes, many test statistics can be constructed to test the equality of quantile regression curves. For example, we can define for $i, j = 1, \dots, k$,

- 1) $T_i = \sup_{0 \leq t \leq 1} |\sqrt{n}\hat{R}_i(t)|$,
- 2) $T_M = \max \left\{ \sup_{0 \leq t \leq 1} |\sqrt{n}\hat{R}_1(t)|, \dots, \sup_{0 \leq t \leq 1} |\sqrt{n}\hat{R}_k(t)| \right\}$.
- 3) $T_{ij} = \sup_{0 \leq t \leq 1} |\sqrt{n}\hat{R}_i(t) - \sqrt{n}\hat{R}_j(t)| \quad i \neq j$.

For a size α test, we can reject H_0 for values of the above statistics exceeding the $(1 - \alpha)$ quantiles of their respective null distribution. To improve finite sample performance of our test statistics we will use a bootstrap method to compute critical values for our tests.

2.1 Bandwidth Selection

Bandwidth selection is an essential part of any nonparametric smoothing procedure. In particular, any testing method that involves nonparametric estimation of functions requires careful attention on bandwidth selection. For local linear nonparametric quantile regression function estimation (the one we have used in our tests), Yu and Jones (1998) have developed a useful rule-of-thumb bandwidth by minimizing the asymptotic mean squared error (AMSE). This bandwidth, h_τ , is calculated as follows:

$$h_\tau = h_{mean} \left\{ \frac{\tau(1-\tau)}{\phi(\Phi^{-1}(\tau))^2} \right\}^{1/5} \quad (6)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density function and the distribution function of the standard normal distribution and h_{mean} is the optimal bandwidth for local linear nonparametric mean regression estimation. This is a computationally easy way to calculate AMSE optimal bandwidths. For example, at $\tau = 0.25$, $h_\tau = h_{mean}(1.13)$ which can be easily obtained using the statistical package R (R Development Core Team, 2010) by $h_\tau = dpill(x, y)(1.13)$. However, we note that this bandwidth selection method is based on several key assumptions such as the conditional density function of the dependent variable can be approximated by normal densities (Yu & Jones, 1998). Bandwidth selection using other criteria such as power optimization is still an open problem in the quantile regression setting.

2.2 Computing Critical Values

Since the exact null distribution is difficult to find analytically, we used a bootstrap method to compute critical values for our test statistics. For quantile regression, the classical bootstrap method do not work since the assumption of zero conditional mean is not true in quantile regression models. To overcome this difficulty Sun (2006) proposed a bootstrap method using a two point wild bootstrap. This method is a modified version of the wild bootstrap of Härdle and Mammen (1993). Since we are comparing our test procedure with that of Sun (2006) we used the bootstrap procedure outlined in Sun (2006) in our simulation study. This method ensures that the τ^{th} conditional quantile of the bootstrap residuals given the data is zero, while the second and the third moments are matched with that of the true residuals.

Let $\hat{\epsilon}_{ij} = Y_{ij} - \hat{g}_\tau(X_{ij})$ denotes the estimated residuals. Here $\hat{g}_\tau(\cdot)$ is the estimated common quantile regression function using (4). Let ϵ_{ij}^b denotes the bootstrap resamples of $\hat{\epsilon}_{ij}$ constructed according to the procedure in Sun (2006). Then we create the bootstrap samples $\{X_{ij}, Y_{ij}^b\}$, where

$$Y_{ij}^b = \hat{g}_\tau(X_{ij}) + \epsilon_{ij}^b, \quad i = 1, \dots, k, \quad j = 1, \dots, n_i.$$

Using these bootstrap samples we then construct B bootstrap replicates of our test statistics. For example, $\{T_{Max,b}\}_{b=1}^B$. Then these bootstrap replicates were used to obtain $(1 - \alpha)$ quantiles of the null distributions of the respective test statistics.

3. Simulation Results

In this section we discuss the results of our simulation study. We generated data from the following heteroscedastic error model

$$Y_{ij} = g_i(X_{ij}) + \sigma(X_{ij})\epsilon_{ij} \quad i = 1, \dots, k, \quad j = 1, \dots, n_i \quad (7)$$

with $X \sim U[0, 1]$ and $\epsilon \sim N(0, 1)$. We conducted our simulations for the case of comparing three groups ($k = 3$). All simulations were repeated 1000 times and we used equal sample sizes for each group. Three sample sizes $n_i = 25, 50, 100$ were considered and five quantiles $\tau = .05, .25, .5, .75, .95$ were examined. It is easy to see that the

τ^{th} conditional quantile function for model (7) is of the form $g_{\tau,t}(\cdot) = g(\cdot) + \sigma(\cdot)Z_{\tau}$ where Z_{τ} is the τ^{th} conditional quantile of ϵ given X . By choosing variety of functions for $g_i(\cdot)$ and $\sigma(\cdot)$ one can create quantile curves having various degrees of smoothness. For example, with $g(x) = x$ and $\sigma(x) = \cos(2\pi x)$, we get the median being linear and more curvature at extreme quantiles as shown in Figure 1.

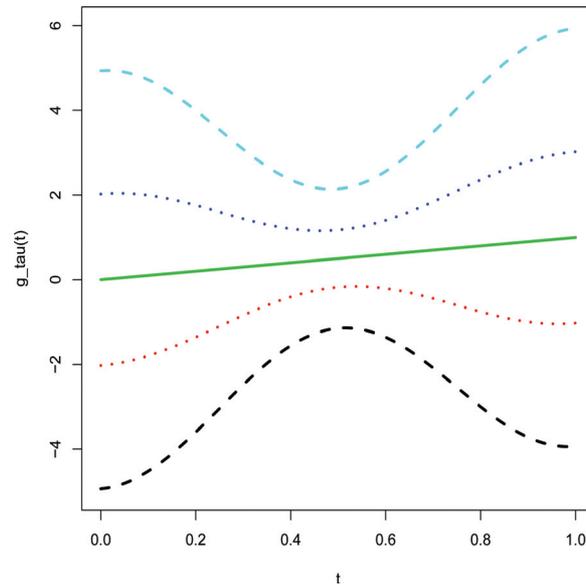


Figure 1. Quantile regression curves correspond to the heteroscedastic model in (7) with $g(t) = t$ and $\sigma(t) = \cos(2\pi t)$. Solid line($\tau = 0.5$), dotted lines($\tau = 0.25, 0.75$) and dashed lines($\tau = 0.05, 0.95$)

We have focused mainly on the test statistic T_{Max} (defined in section 2) in our simulation study as it is the natural choice when comparing more than two groups. However, for comparison purposes we present results for two other statistics, namely, T_1 and T_{12} . The simulation study consists of three parts:

- 1) Investigate the Type 1 error and power properties of our test statistics in comparison to Sun (2006).
- 2) Investigate how our test procedure can be used as an alternative to nonparametric mean regression procedures in comparing groups, especially in the presence of extreme observations. We compared our approach at $\tau = 0.5$ (conditional median) with the nonparametric mean regression test proposed by Neumeyer and Dette (2003). Both Type 1 error and power of the two methods were studied with 100% normal errors and with a mixture of 80% normal and 20% Cauchy errors.
- 3) Examine local power properties of our test statistics in comparison to Sun (2006).

3.1 Analysis of Type I Error and Power

In this section we compare the Type 1 error of our test statistics with Sun (2006). We present two cases: Case 1 represents a situation of heteroscedastic error variation where the variability increases linearly; Case 2 represents a situation where the variability increases non-linearly. The critical values are computed for a α -level test with $\alpha = 0.05$. As shown in Table 1, the Type I errors of our test statistic T_{Max} underestimate the level of the test for small sample sizes but gets to the nominal 5% level as the sample size increases. In comparison, the test statistic proposed in Sun (2006) (denoted by SUN), underestimate the level at upper quantiles even for large sample sizes. This pattern is more prominent in Case 2, where the error variance is nonlinear.

Table 1. Type I error based on 1000 simulated samples

Functions	n	τ	T_{Max}	T_1	T_{12}	SUN
Case 1						
$g_1 = g_2 = g_3 = \exp(t)$ $\sigma_1(t) = \sigma_2(t) = \sigma_3(t) = 1 + t$	25	0.05	0.045	0.033	0.034	0.039
		0.25	0.042	0.044	0.057	0.042
		0.50	0.044	0.054	0.062	0.048
		0.75	0.038	0.041	0.058	0.040
		0.95	0.033	0.026	0.028	0.042
50	0.05	0.045	0.042	0.042	0.055	
	0.25	0.051	0.050	0.069	0.037	
	0.50	0.036	0.034	0.048	0.036	
	0.75	0.044	0.039	0.047	0.030	
	0.95	0.040	0.017	0.040	0.042	
100	0.05	0.050	0.050	0.043	0.047	
	0.25	0.044	0.052	0.064	0.041	
	0.50	0.032	0.043	0.053	0.032	
	0.75	0.041	0.036	0.057	0.027	
	0.95	0.050	0.043	0.051	0.032	
Case 2						
$g_1 = g_2 = g_3 = \exp(t)$ $\sigma_1(t) = \sigma_2(t) = \sigma_3(t) = 1 + t + \cos(2\pi t)$	25	0.05	0.060	0.043	0.042	0.055
		0.25	0.049	0.037	0.062	0.037
		0.50	0.032	0.044	0.056	0.048
		0.75	0.030	0.037	0.051	0.027
		0.95	0.032	0.011	0.024	0.041
50	0.05	0.054	0.031	0.050	0.046	
	0.25	0.040	0.033	0.078	0.027	
	0.50	0.030	0.033	0.042	0.030	
	0.75	0.023	0.022	0.065	0.016	
	0.95	0.017	0.009	0.043	0.048	
100	0.05	0.033	0.030	0.063	0.024	
	0.25	0.023	0.026	0.067	0.021	
	0.50	0.033	0.040	0.055	0.022	
	0.75	0.012	0.018	0.052	0.008	
	0.95	0.035	0.037	0.067	0.027	

Next we examine the power of our test statistics and compare it with the test statistic proposed in Sun (2006). To minimize the complexity of the experiment we used a constant variance function for this example. Table 2 summarizes the results. Both our test statistic and Sun's method have reasonable power even with small sample sizes as shown in Table 2. As sample size grows both methods approach 100% power in for mid quantiles. At extreme quantiles, our T_{Max} statistics performs better compared to Sun's procedure. For example, at the 95th percentile, the difference in power is about 14% in favor of our T_{Max} statistics.

Table 2. Empirical Power based on 1000 simulated samples

Functions	n	τ	T_{Max}	T_1	T_{12}	SUN
$g_1(t) = t$ $g_2(t) = e^t$ $g_3(t) = \sin(2\pi t)$ $\sigma_1(t) = \sigma_2(t) = \sigma_3(t) = 0.1$	25	0.05	0.043	0.014	0.010	0.144
		0.25	0.783	0.204	0.496	0.881
		0.50	0.975	0.398	0.948	0.979
		0.75	0.937	0.630	0.979	0.911
		0.95	0.078	0.000	0.101	0.145
	50	0.05	0.153	0.025	0.022	0.142
		0.25	0.996	0.396	0.856	1.000
		0.50	1.000	0.756	1.000	1.000
		0.75	1.000	0.963	1.000	1.000
		0.95	0.371	0.002	0.574	0.289
	100	0.05	0.548	0.037	0.058	0.473
		0.25	1.000	0.704	0.993	1.000
		0.50	1.000	0.972	1.000	1.000
		0.75	1.000	0.999	1.000	1.000
		0.95	0.847	0.683	0.956	0.696

3.2 Comparison With Nonparametric Mean Regression

As described in section 1, the mean regression approach, although quite popular in comparing groups, will have severe drawbacks in the presence of extreme observations. Therefore, we examine how our test procedure can be used as an alternative for mean regression testing methods when the data is contaminated with extreme values. To this end, we compare our testing procedure at $\tau = 0.5$ (conditional median) with the non parametric mean regression test proposed by Neumeyer and Dette (2003) under two error structures. We used model (7) with $\epsilon \sim N(0, 1)$ and $\epsilon \sim 0.8N(0, 1) + 0.2Cauchy$. The first error structure mimics the classical regression setup and the second error structure creates outliers by having 20% of the errors come from the standard Cauchy distribution. The choice of the contamination error distribution as Cauchy was motivated by its well-known heavy tail structure. Although any other heavy tail distribution could also be used, we feel that the Cauchy case will serve as a “lower bound” for the power properties of our tests. The weights of the mixture (0.8 and 0.2) was chosen to provide a reasonable balance between the level of contamination and the data. In order to be compatible with the simulation study of Neumeyer and Dette(2003) we used the same functions in their study and focused on comparing two groups ($k = 2$ case) with $\sigma(t) = 0.5$.

Table 3 summarizes the results of the experiment with normal errors. Both quantile regression approach (T_{Max} and SUN) and mean regression approach (ND) performed quite well under the normal error structure. The Type 1 error of ND is closer to the 0.05 nominal level than the other two procedures. With linear alternatives (3rd and 4th examples in Table 3), T_{Max} performed as good as the mean regression method in large samples. The power of the quantile regression methods is slightly lower than the mean regression method when dealing with high frequency alternatives (last two examples in Table 3). This is not surprising because the mean regression approach is designed under the normal error assumption with its power properties optimized based on moment conditions of normal errors.

Next we examine the Type 1 error and power properties of these methods under a mixture error structure ($\epsilon \sim 0.8N(0, 1) + 0.2Cauchy$). We used the same examples as in Table 3. The results are summarized in Table 4. In contrast, the values in Table 4 clearly shows that the mean regression method breaks down, especially the power of detecting the difference in the two regression functions. For example, in testing $g_1(t) = \sin(2\pi t)$, $g_2(t) = \sin(2\pi t) + t$, with a sample size $n = 100$, the power of the mean regression method drops dramatically from 100% (with normal errors) to 28% (with a mixture of normal and Cauchy errors). In comparison, the power of the proposed method T_{Max} in particular, remains very stable (97% to 93%).

Table 3. Type I error and power with 100% normal errors

<i>Functions</i>	<i>n</i>	<i>T_{Max}</i>	<i>T₁</i>	<i>T₁₂</i>	<i>SUN</i>	<i>ND</i>
Type I error						
$g_1(t) = g_2(t) = e^t$	25	0.032	0.036	0.059	0.028	0.053
	50	0.028	0.034	0.053	0.022	0.067
	100	0.035	0.044	0.062	0.033	0.059
$g_1(t) = g_2(t) = \sin(2\pi t)$	25	0.029	0.029	0.040	0.014	0.055
	50	0.028	0.029	0.060	0.003	0.056
	100	0.013	0.019	0.076	0.001	0.040
Power						
$g_1(t) = e^t, g_2(t) = g_1(t) + t$	25	0.502	0.557	0.625	0.500	0.915
	50	0.869	0.902	0.942	0.852	0.995
	100	0.997	0.990	0.999	0.999	1.000
$g_1(t) = \sin(2\pi t), g_2(t) = g_1(t) + t$	25	0.435	0.433	0.592	0.350	0.754
	50	0.759	0.721	0.903	0.558	0.970
	100	0.974	0.952	0.996	0.814	1.000
$g_1(t) = e^t, g_2(t) = g_1(t) + \sin(2\pi t)$	25	0.156	0.134	0.264	0.679	0.385
	50	0.473	0.321	0.623	0.946	0.840
	100	0.928	0.778	0.970	1.000	0.997
$g_1(t) = \sin(2\pi t), g_2(t) = 2\sin(2\pi t)$	25	0.212	0.038	0.203	0.433	0.211
	50	0.533	0.042	0.526	0.652	0.536
	100	0.878	0.099	0.912	0.880	0.951

Table 4. Type I error and power with 80% normal and 20% Cauchy errors

<i>Functions</i>	<i>n</i>	<i>T_{Max}</i>	<i>T₁</i>	<i>T₁₂</i>	<i>SUN</i>	<i>ND</i>
Type I error						
$g_1(t) = g_2(t) = e^t$	25	0.030	0.033	0.045	0.041	0.033
	50	0.035	0.030	0.053	0.040	0.039
	100	0.032	0.042	0.063	0.021	0.027
$g_1(t) = g_2(t) = \sin(2\pi t)$	25	0.041	0.037	0.050	0.021	0.030
	50	0.024	0.041	0.056	0.009	0.033
	100	0.009	0.020	0.057	0.000	0.031
Power						
$g_1(t) = e^t, g_2(t) = g_1(t) + t$	25	0.385	0.446	0.505	0.345	0.261
	50	0.753	0.791	0.848	0.697	0.280
	100	0.971	0.979	0.989	0.961	0.271
$g_1(t) = \sin(2\pi t), g_2(t) = g_1(t) + t$	25	0.343	0.362	0.454	0.241	0.229
	50	0.651	0.617	0.802	0.429	0.272
	100	0.927	0.889	0.990	0.647	0.279
$g_1(t) = e^t, g_2(t) = g_1(t) + \sin(2\pi t)$	25	0.109	0.109	0.198	0.488	0.084
	50	0.354	0.243	0.473	0.840	0.112
	100	0.807	0.597	0.881	0.994	0.135
$g_1(t) = \sin(2\pi t), g_2(t) = 2\sin(2\pi t)$	25	0.136	0.043	0.177	0.284	0.061
	50	0.403	0.041	0.411	0.493	0.092
	100	0.769	0.074	0.811	0.748	0.120

3.3 Local Power Properties

The goal of this experiment is to provide evidence that our proposed method has a faster rate of convergence than the test statistic proposed by Sun(2006). We examine the power properties of our test procedure with local alternatives converging to the null at $(\sqrt{n})^{-1}$ rate. We compare our results to that of Sun (2006) . We restricted our simulation to the $k = 2$ case to save computational time. We used model (7) with $g_1(t) = e^t, g_2(t) = g_1(t) + \delta/\sqrt{n}$ and $\sigma(t) = 0.1$ (homoscedastic errors). Here δ is a constant that controls the separation of the two functions and we chose $\delta = 0.3$ in our simulation. The sample sizes for each group were chosen large enough to make the testing problem difficult. Three sample sizes $n_1 = n_2 = 500, 2500, 5000$ were considered and three quantiles $\tau = .05, .50, .75$ were examined. All simulations were repeated 1000 times and the results are given in Table 5.

Table 5. Comparison of power with local alternatives with $\alpha = 0.05$

τ	<i>N</i>	<i>T_{Max}</i>	<i>T₁</i>	<i>T₁₂</i>	<i>SUN</i>
0.05	1000	0.244	0.285	0.285	0.072
	5000	0.258	0.287	0.292	0.050
	10000	0.273	0.308	0.312	0.050
0.50	1000	0.431	0.532	0.635	0.153
	5000	0.301	0.425	0.662	0.077
	10000	0.242	0.358	0.669	0.032
0.75	1000	0.417	0.519	0.588	0.134
	5000	0.299	0.415	0.584	0.073
	10000	0.266	0.384	0.618	0.050

As shown in Table 5, our test statistics have much better local power properties compared to Sun (2006). Our tests remain well above the nominal $\alpha = 5\%$ level whereas the power of Sun's test converge down to α . As suggested by one of the referees, in order to get a more general conclusion we repeated the experiment under another setting. We used different functions: $g_1(t) = \sin(2\pi t)$, $g_2(t) = g_1(t) + \delta/\sqrt{n}$ and $\sigma(t) = 0.1$ and used $\alpha = 0.01$. As above, all simulations were repeated 1000 times and the results are given in Table 6.

Table 6. Comparison of power with local alternatives with $\alpha = 0.01$

τ	N	T_{Max}	T_1	T_{12}	SUN
0.05	1000	0.081	0.091	0.096	0.021
	5000	0.150	0.093	0.084	0.027
	10000	0.187	0.102	0.099	0.019
0.50	1000	0.098	0.211	0.246	0.044
	5000	0.145	0.204	0.275	0.031
	10000	0.196	0.212	0.284	0.028
0.75	1000	0.126	0.131	0.261	0.031
	5000	0.131	0.146	0.215	0.026
	10000	0.173	0.182	0.278	0.020

As expected, when the α value is low (0.01), the power goes down in both methods, but our method exhibits higher power than the nominal α level. Since the alternative hypotheses that we used in these two experiments are converging at $(\sqrt{n})^{-1}$ rate, these results clearly demonstrate that our test statistics have a faster convergence rate than Sun (2006).

4. Data Analyses

We now present two applications of our test procedure to real data:

- 1) Analysis of maximum wind speed of tropical cyclones in the Atlantic basin.
- 2) Analysis of managerial expenses of Japanese chemical firms.

4.1 Example 1: Tropical Cyclones in the Atlantic Basin

As an application of our proposed methodology we consider a data analysis on Atlantic basin tropical cyclones in the HURDAT data set, which is available on the National Oceanic Atmospheric Administration's website (www.aoml.noaa.gov/hrd/hurdat). The current data contain information on storms between 1851 and 2010. A changepoint analysis of the data through 2008 conducted in Robbins, Lund, Gallagher, and Lu (2011) found changepoints in several measured variables in the early to mid 1960's. Satellites were fully implemented for cyclone surveying in the mid 1960 s, and their introduction also likely improved the accuracy of measured wind speeds. For the above reasons we consider only data from 1968-2010. A changepoint test in Robbins, Lund, Gallagher, and Lu (2011) indicated a significant change in the frequency of storms around 1995. We therefore divide storms into two groups 1968-1994 ($n_1 = 258$) and 1995-2010 ($n_2 = 236$).

There has been much interest in whether storms have become stronger in recent years. There are many measures of storm strength considered in the literature (Emanuel, 2005; Briggs, 2008). Here we simply measure storm intensity as the maximum observed wind speed of the cyclone. We investigate the behavior of the 0.9 quantile of maximum windspeed as a function of position within the hurricane season. In particular let $X = (\text{storm number in } i^{\text{th}} \text{ year})/(\text{total number of storms in year } i)$, e.g., the fourth storm in a year with 11 storms would have an associated X value of 4/11. Figure 2 contains a plot of the data. The overlay of the nonparametric estimates of the quantile curves indicates that the basic shape of the quantile curve is the same for the two groups, with the strongest storms tending to occur near the middle of the season and relatively fewer strong storms occurring at the beginning of the season. Applying the test statistic T_{Max} to the data indicates a significant difference between the two curves at the 0.05 significance level (test stat = 6.813, cutoff = 6.416). Since the curve for the more recent storms (1995-2010) is systematically above that of the storms from 1968-1994, this provides some evidence that the .9 quantile of maximum wind speed has increased in recent years. We note that the procedure of Sun (2006) did not detect a significant difference between the two curves (test stat = 0.095, cutoff = 0.166). Furthermore, the nonparametric

mean regression test of Neumeyer and Dette (2003) also did not detect a difference between the two groups. Our test provides different conclusions than either of those techniques.

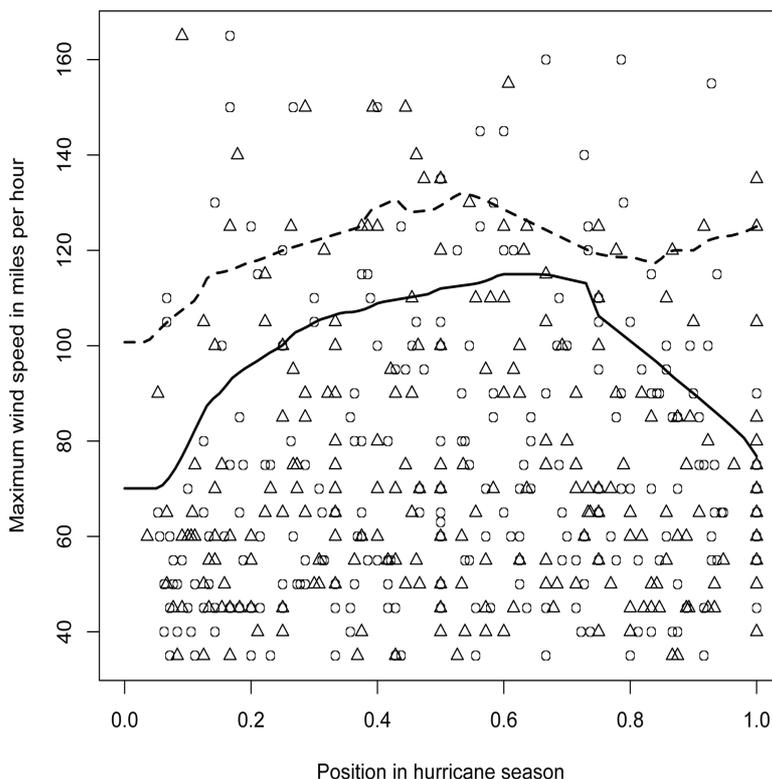


Figure 2. Plot of hurricane data from 1968-2010. Circles represent before 1995, triangles represent after 1995 data. 90th percentile curves: — before 1995, - - - after 1995

4.2 Example 2: Managerial Expenses of Japanese Chemical Firms

We analyzed the Japanese chemical industry data (Yafeh & Yosha, 2003) which is publicly available at the *Economic Journal* website at <http://www.res.org.uk>. Data consists of various economic factors collected on 185 chemical firms listed in the Japanese stock market. The dependent variable is a measure of expenses on managerial private benefits. The study was focused on investigating the hypothesis that “concentrated shareholding is associated with lower expenditure on managerial private benefits”. Ownership concentration of the firm is measured by the variable TOPTEN, which gives the percentage of ownership that belongs to the top 10 shareholders. Our investigation was to check whether this association is different among firms that has high/low leverage (the ratio of debt to debt plus equity). We divided our data into two groups based on the leverage measurements of the firms. Firms with leverage values above the sample median are categorized in to ‘high’ and the rest are into ‘low’ groups.

Figure 3 depicts how managerial expenses (Y) change with ownership concentration (TOPTEN) in the two groups.

As shown in Figure 3, there is a negative association between expenditure on managerial private benefits and the ownership concentration in both groups. The nonparametric mean regression estimates (solid and dotted lines) suggest that the two groups are different. A hypothesis test (Neumeyer & Dette, 2003) revealed that the two mean regression curves are in fact significantly different at .05 level. We then used our test procedure and the one proposed by Sun (2006) at $\tau = 0.5$ to check whether the conditional medians of the two groups are the same or not. All of our test statistics were able to capture a difference between the two conditional medians. However, the test proposed in Sun (2006) did not reject the null hypothesis of equality of conditional median curves. We used the bandwidth selection methods which we adopted in our simulations. For computing bootstrap critical values, we used $B = 1000$ repetitions to increase accuracy. Table 7 summarizes the results of our testing at several conditional quantiles. This analysis suggests that the differences between two conditional distributions exists at mid to upper quantiles. No differences was observed at quantiles below the median. It provides a more elaborate picture of the actual differences in the two groups compared to the nonparametric mean regression approach of Neumeyer and

Dette (2003).

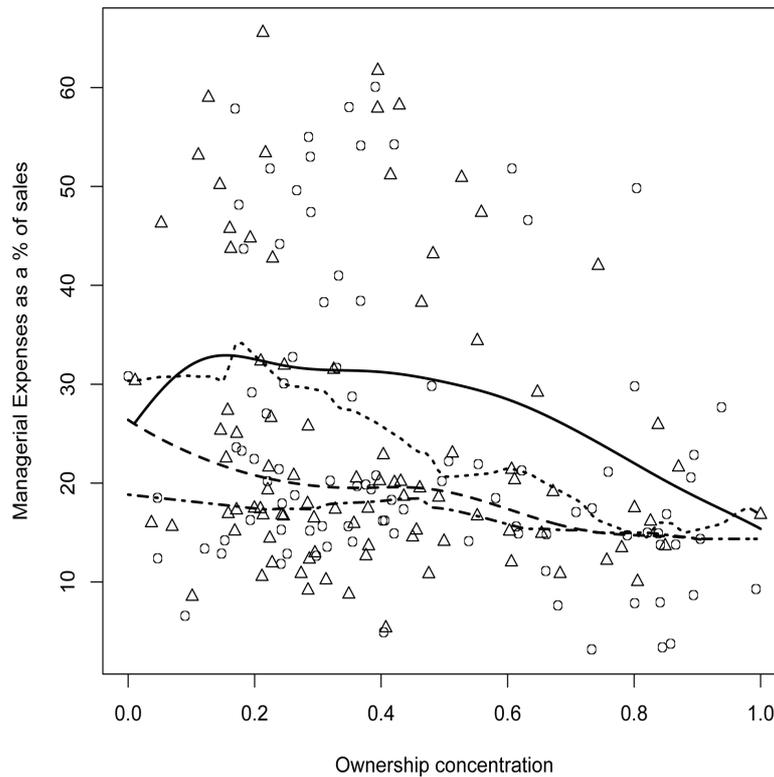


Figure 3. Plot of managerial expenses vs shareholder concentration of Japanese Chemical firms. Circles-low leverage, triangles-high leverage. The plotted lines are the nonparametric mean and median regression estimates for the two groups. — mean reg (low), ··· mean reg (high), - - - q-reg (low), · · · q-reg (high)

Table 7. Test results of the Japanese chemical industry data $\alpha = 0.05$

τ	T_{Max}	T_1	T_{12}	SUN
0.05	N	N	N	N
0.25	N	N	N	N
0.50	R	R	R	N
0.75	R	R	R	R
0.95	R	R	R	R

Note: R-Rejected , N-did not reject.

5. Conclusion

We have proposed a robust nonparametric test of detecting the differences in conditional quantile functions for multiple populations. The proposed method demonstrates a desirable \sqrt{n} property in contrast to the slower rates of Sun (2006). Bandwidth selection for nonparametric quantile estimation and for tests based on such estimators is an open problem. In this article we have used a plug in rule that is readily available. A detailed study of bandwidths that optimize power while preserving size is highly warranted.

References

Bondell, H. D., Reich, B. J., & Wang, H. (2010). Non-crossing quantile regression curve estimation. *Biometrika*, 70(3), 609-627.

Briggs, W. M. (2008). On the changes in the number and intensity of north Atlantic tropical cyclones. *Journal of Climate*, 21, 1387-1402. <http://dx.doi.org/10.1175/2007JCLI1871.1>

Delgado, M. A. (1993). Testing the equality of nonparametric regression curves. *Statistics and Probability Letters*,

17, 199-204.

- Dette, H., & Volgushev, S. (2008). Non-crossing non-parametric estimates of quantile curves. *Journal of the Royal Statistical Society, Ser. B*, 70(3), 609-627.
- Emanuel, K. (2005). Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, 436, 686-689. <http://dx.doi.org/10.1038/nature03906>
- Härdle, W., & Mammen, E. (1993). Comparing Nonparametric Versus Parametric Regression Fits. *The Annals of Statistics*, 21, 1926-1947.
- Hall, P., & Hart, J. D. (1990). Bootstrap Test for Difference Between Means in Nonparametric Regression. *Journal of the American Statistical Association*, 85, 1039-1049.
- King, E. C., Hart, J. D., & Wehrly, T. E. (1991). Testing the equality of two regression curves using linear smoothers. *Statistics and Probability Letters*, 12, 239-247.
- Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica*, 46, 33-50.
- Kulasekera, K. B. (1995). Comparison of Regression Curves Using Quasi-Residuals. *Journal of the American Statistical Association*, 90, 1085-1093.
- Neumeier, N., & Dette, H. (2003). Nonparametric Comparison of Regression Curves: An Empirical Process Approach. *The Annals of Statistics*, 31, 880-920.
- R Development Core Team. (2010). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>
- Robbins, M., Lund, R., Gallagher, C., & Lu, X. (2011). A change-point analysis of Atlantic tropical cyclones. *Journal of the American Statistical Association*, 106, 89-99. <http://dx.doi.org/10.1198/jasa.2011.ap10023>
- Ruppert, D., Sheather, S. J., & Wand, M. P. (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association*, 90, 1257-1270.
- Sun, Y. (2006). A Consistent Nonparametric Equality Test of Conditional Quantile Functions. *Econometric Theory*, 22, 614-632.
- Takeuchi, Q. V., Sears, T. D., & Smola, A. J. (2006). Nonparametric Quantile Estimation. *Journal of Machine Learning Research*, 7, 1231-1264.
- Yafeh, Y., & Yosha, O. (2003). Large shareholders and banks: Who monitors and how? *Economic Journal*, 113, 128-146.
- Yu, K., & Jones, M. C. (2003). Local Linear Quantile Regression. *Journal of the American Statistical Association*, 93, 228-289.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).