


11-1-2007

# An Omnibus Test When Using a Regression Estimator With Multiple Predictors

Rand R. Wilcox

University of Southern California, [rwilcox@usc.edu](mailto:rwilcox@usc.edu)

Follow this and additional works at: <http://digitalcommons.wayne.edu/jmasm>

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

## Recommended Citation

Wilcox, Rand R. (2007) "An Omnibus Test When Using a Regression Estimator With Multiple Predictors," *Journal of Modern Applied Statistical Methods*: Vol. 6 : Iss. 2 , Article 3.

DOI: 10.22237/jmasm/1193889720

Available at: <http://digitalcommons.wayne.edu/jmasm/vol6/iss2/3>

This Invited Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.

## An Omnibus Test When Using A Regression Estimator With Multiple Predictors



Rand R. Wilcox  
University of South Carolina

---

In quantile regression, the goal is to estimate the  $\gamma$  quantile of  $Y$  given values for  $p$  predictors. Methods for making inferences about the individual slope parameters have been proposed, some of which have been found to perform very well in simulations. But for an omnibus test that all slope parameters are zero, it appears that little is known about how best to proceed. For the special case  $\gamma=.5$ , a drop-in-dispersion test has been recommended, but it requires a large sample size to control the probability of a Type I error and it assumes that the usual error term is homoscedastic. The article suggests an alternative method that performs well in simulations, it allows heteroscedasticity, and it can be used when  $\gamma \neq .5$ .

Key words: Robust regression, tests of independence, bootstrap methods.

---

### Introduction

Consider the random variables  $X_1, \dots, X_p, Y$  having some unknown  $(p+1)$ -variate distribution and let  $Y_\gamma$  be the conditional  $\gamma$  quantile of  $Y$  given  $X_1, \dots, X_p$ . When using the Koenker and Bassett (1978) quantile regression method, the goal is to estimate  $Y_\gamma$  assuming that

$$Y_\gamma = \alpha_\gamma + \beta_{1\gamma}X_1 + \dots + \beta_{p\gamma}X_p \quad (1)$$

---

Rand R. Wilcox (rwilcox@usc.edu) is Professor of Psychology. He is the author of seven textbooks on statistics, the most recent of which is *Introduction to Robust Estimation and Hypothesis Testing* (2005, 2<sup>nd</sup> ed., San Diego, CA: Academic Press).

where the unknown parameters  $\beta_{1\gamma}, \dots, \beta_{p\gamma}$  and  $\alpha_\gamma$  are estimated based on the random sample  $(X_{i1}, \dots, X_{ip}, Y_i)$ ,  $i = 1, \dots, n$ . The special case  $\gamma=.5$  corresponds to what is called the least absolute value regression estimator, meaning that the estimates of the parameters are chosen so as to minimize the sum of the absolute values of the residuals. This special case predates ordinary least squares by about a half century and offers protection against the deleterious effects of outliers among the  $Y$  values. As is probably evident, choices for  $\gamma$  other than  $.5$  can be revealing and help add perspective on the association among the variables under study.

As a simple example, consider data from a study conducted by Williams, Stanchina,

Table. 1 Values for  $d_0$  and  $d_1$ 

p	$d_0$				$d_1$			
	$\alpha=.1$	$\alpha=.05$	$\alpha=.025$	$\alpha=.01$	$\alpha=.1$	$\alpha=.05$	$\alpha=.025$	$\alpha=.01$
2	.2179	.1203	.0588	.0430	-.00196	-.00117	-.00056	-.00055
3	.2814	.1840	.1143	.0364	-.00300	-.00223	-.00149	-.00044
4	.4478	.3356	.2624	.1546	-.00580	-.00476	-.00396	-.00240
5	.6373	.4250	.3097	.1590	-.00896	-.00630	-.00474	-.00248
6	.7699	.5648	.4111	.2734	-.01120	-.00858	-.00640	-.00439

Bezdjian, Skrok, Raine and Baker (2005). A portion of the study dealt with the association between a so-called Q score resulting from the Porteus maze test, which is used to evaluate intelligence and executive functioning, and how this Q score is related to a measure of delinquency. Figure 1 shows a scatterplot of the data. The sample size is  $n=943$ . Also shown are the regression lines corresponding to  $\gamma=.5$ ,  $.8$  and  $.9$ . As is evident, based on the typical response, as measured by the median or even the  $.8$  quantile, there is little or no indication of an association. (The p-value when  $\gamma=.8$  is approximately  $.36$ .) But for  $\gamma=.9$ , the regression line has a positive slope that is significantly different from zero at the  $.05$  level. (For another recent illustration of the practical value of quantile regression methods, see Angrist, Chernozhukov & Fernandez-Val, 2006.)

The goal in this article is to suggest and study a method for testing

$$H_0 : \beta_{1\gamma} = \dots = \beta_{p\gamma} = 0. \quad (2)$$

For the related problem of testing

$$H_0 : \beta_j = 0$$

for each  $j$  ( $j=1, \dots, p$ ), there is a well-known method that appears to perform relatively well in simulations (Koenker, 1994, cf. Koenker & Xiao, 2002, cf. Koenker & Machado, 1999). But when  $\gamma$  differs from  $.5$ , it seems that there are no results or even suggested methods for testing (2).

For the special case  $\gamma=.5$ , Birkes and Dodge (1993) suggest testing (2) using a drop in dispersion method. They note that the method requires a relatively large sample size, but they do not specify just how large the sample size must be to achieve reasonably accurate control over the probability of a Type I error. When testing at the  $.05$  level, Bradley (1978) suggests that at a minimum, the actual Type I error probability should be between  $.025$  and  $.075$ . When examining the drop in dispersion method (in the simulations described in section 3), it was found that to achieve Bradley's criterion, a sample size of  $n=100$  is required, even under normality. Another concern is that the method assumes a homoscedastic error term. So one goal here is search for a method that gives better results when the sample size is small and another goal is to suggest a method that might be used when the error term is heteroscedastic.

Yet another approach to testing (2) is to use the percentile bootstrap method stemming from results in Liu and Singh (1997). When working with various robust estimators, this approach appears to perform quite well, even with fairly small sample sizes and when there is heteroscedasticity (e.g., Wilcox, 2005). However, this approach was found to be unsatisfactory in the simulations considered here, so it was abandoned.

### Methodology

The Koenker and Bassett (1978) quantile regression method arises as follows.

For some  $\gamma, 0 < \gamma < 1$ , let

$$\rho_\gamma(u) = u(\gamma - I_{u < 0})$$

where the indicator function  $I_{u < 0} = 1$  if  $u < 0$ ; otherwise  $I_{u < 0} = 0$ . Assuming that the  $\gamma$  quantile of  $Y$ , given  $X$ , is given by (1), the Koenker-Bassett quantile regression method estimates the unknown parameters  $\beta_{1\gamma}, \dots, \beta_{p\gamma}$  and  $\alpha_\gamma$  with the values  $b_{1\gamma}, \dots, b_{p\gamma}$  and  $a_\gamma$ , respectively, that minimize

$$\sum \rho_\gamma(r_i), \tag{3}$$

where  $r_i = Y_i - b_{1\gamma}X_{i1} - \dots - b_{p\gamma}X_{ip} - a_\gamma$  are the residuals. Here, the values that minimize (3) were determined with the function `rq` that is included in the robust library that comes with the software S-PLUS.

The proposed method for dealing with small sample sizes stems in part from the classic generalized  $T^2$  statistic used to test the hypothesis that a multivariate normal distribution has a mean vector of zero (e.g., Anderson, 1958, chapter 5). One difficulty here is getting an estimate of the appropriate covariance matrix, and the strategy is to use a bootstrap estimate. (For general results on bootstrap estimates of the standard error, see Buchinsky, 1991; Hahn, 1995.) Results for the special case  $p=1$ , reported by Koenker (1994), suggest that this approach will result in an actual Type I error probability that can be substantially less than the nominal level, and this was found to be the case for  $n < 60$ . However, a simple adjustment is found that corrects this problem in the simulations to be described.

Let  $(X_{i1}^*, \dots, X_{ip}^*, Y_i^*)$ ,  $i = 1, \dots, n$ , be a bootstrap sample obtained by randomly sampling, with replacement,  $n$  vectors of observations from  $X_{i1}, \dots, X_{ip}, Y_i$ . Given  $\gamma$ , label the resulting estimate of the slopes  $b_k^*$ ,  $k = 1, \dots, p$ . Repeat this process  $B$  times yielding  $b_{1k}^*, \dots, b_{pk}^*$ . Then from basic principles,

an estimate of the variances and covariances associated with  $b_{1\gamma}, \dots, b_{p\gamma}$  is

$$S = \frac{1}{B-1} \sum_{c=1}^B (b_c^* - \bar{b}^*)^2,$$

where  $b_c^* = (b_{c1}^*, \dots, b_{cp}^*)$ , and  $\bar{b}^* = \sum b_{ck}^* / B$ . Then, proceeding in an obvious fashion, the test statistic used here is

$$T^2 = n\bar{b}'S^{-1}\bar{b}.$$

Again from basic principles, a natural strategy is to reject if

$$T^2 \geq \frac{n-1}{n-p} f_{p, n-p},$$

where  $f_{p, n-p}$  is the  $1-\alpha$  quantile of an F distribution with  $p$  and  $n-p$  degrees of freedom. But as previously indicated, preliminary simulations indicated that the actual probability of Type I error is less than the nominal level when the sample size is small. For example, when  $\gamma=.5, p=2, n=20, \alpha=.05$ , and if  $X_1$  and  $X_2$  have a bivariate normal distribution with correlation  $\rho=0$ , the actual Type I error probability was estimated to be .026. Increasing  $p$  to 6, the estimate is now .001. Very similar results were obtained when  $\gamma=.8$ . But in all cases considered, with  $n=60$ , the actual probability of a Type I error was estimated to be reasonably close to .05.

The results just described suggest the following modification when  $n < 60$ .

Temporarily assume that the error term is homoscedastic and has a normal distribution. The strategy is to determine an adjusted  $p$ -value,  $p_a$ , so that for  $n=20$ , the actual Type I error probability will be approximately  $\alpha$  if the null hypothesis is rejected  $\hat{p} \leq p_a$  whenever the observed  $p$ -value (based on  $T^2$ ) is less than or equal to  $p_a$ . (In essence, use Gosset's strategy when dealing with the problem of making inferences about means.) For sample sizes between 20 and 60, interpolation is used to

Table 2. Some properties of the g-and-h distributions

g	h	$\kappa_1$	$\kappa_2$
0	0	0	3
0	0.2	0	21.46
0.2	0	0.61	3.68
0.2	0.2	2.81	155.98

determine  $p_a$ . First consider  $\gamma=.5$ . For  $\alpha=.1$ , .05, .025 and .01, simulations indicate that the adjusted p-value is given by  $p_a = d_1 n + d_0$ , where  $d_1$  and  $d_0$  are given in Table 1. That is, letting  $\hat{p}$  be the p-value based on  $T^2$ , and assuming that  $(n-p)T^2/(n-1)$  has an F distribution with p and n-p degrees of freedom, reject if  $\hat{p} \leq p_a$ . Additional simulations indicate that this adjustment continues to perform reasonably well when  $\gamma=.8$ , provided B=200 is used, as will be seen.

#### A Simulation Study

Simulations were used to study the small-sample properties of the method just described. The distribution for X was taken to be multivariate normal with common correlation  $\rho$ , and the distribution for Y was taken to be one of four g-and-h distributions (Hoaglin, 1985), which contains the standard normal distribution as a special case. If Z has a standard normal distribution, then

$$Y = \exp\left(\frac{gZ-1}{g}\right) \exp(hX^2/2)$$

if  $g > 0$

$$Y = Z \exp(hZ^2/2)$$

if  $g=0$  has a g-and-h distribution where g and h are parameters that determine the first four moments. The four distributions used here were the standard normal ( $g=h=0.0$ ), a symmetric heavy-tailed distribution ( $h=0.2$ ,  $g=0.0$ ), an asymmetric distribution with relatively light tails ( $h=0.0$ ,  $g=0.2$ ), and an asymmetric distribution

with heavy tails ( $g=h=0.0$ ). Table 2 shows the skewness ( $\kappa_1$ ) and kurtosis ( $\kappa_2$ ) for each distribution considered. Additional properties of the g-and-h distribution are summarized by Hoaglin (1985). The two choices for  $\rho$  were 0 and .8. It was found that altering  $\rho$  had no effect on the simulation results, so for brevity, only results for  $\rho=0$  are reported.

To get some indication of the effects of heteroscedasticity, data were also generated according to the model

$$Y = \lambda(X_1)\epsilon$$

for some specified function  $\lambda$ , where  $\epsilon$  is independent of  $X_1$  and  $\epsilon$  has one of the g-and-h distributions already described. Of course  $\lambda(X_1)=1$  corresponds to homoscedasticity. The other two choices were  $\lambda(X_1)=|X_1|+1$  and  $\lambda(X_1)=1/(|X_1|+1)$ . For convenience, these three choices will be called variance patterns VP1, VP2 and VP3, respectively. Note that for all three patterns, the slope remains zero even when  $\gamma \neq .5$ .

Table 3 shows the estimated probability of a Type I error when testing at the .05 level with  $n=20$ ,  $\gamma=.5$  and .8, and  $p=2$  and 6. For the moment, B=100 is used. It will be seen that generally this suffices, in terms of controlling the probability of a Type I error, but in some cases, B=200 is required. The estimated Type I error probabilities are based on 1,000 replications.

From Robey and Barcikowski (1992), 1,000 replications is sufficient from a power point of view. More specifically, if one tests the hypothesis that the actual Type I error rate is .05,

Table 3: Estimated probability of a Type I error, n=20,  $\alpha = .05$ , B=100

		p=2			p=6		
g	h	VP1	VP2	VP3	VP1	VP2	VP3
$\gamma = .5$							
0.0	0.0	.043	.067	.020	.044	.027	.036
0.0	0.2	.026	.032	.021	.018	.023	.018
0.2	0.0	.037	.052	.023	.034	.048	.028
0.2	0.2	.030	.035	.020	.017	.024	.016
$\gamma = .8$							
0.0	0.0	.049	.029	.076	.048	.046	.032
0.0	0.2	.044	.069	.031	.034	.030	.031
0.2	0.0	.050	.089	.035	.056	.052	.036
0.2	0.2	.050	.078	.030	.041	.041	.036

and if one wants power to be .9 when testing at the .05 level and the true  $\alpha$  value differs from .05 by .025, then 976 replications are required. As is evident, all indications are that reasonable control over the probability of a Type I error is obtained in nearly all of the situations considered. The main exception is when p=2,  $\gamma = .8$  and sampling is from a light-tailed distribution (h=0), in which case, for variance pattern VP2, the estimated probability of a Type I error can exceed .075. The least satisfactory result was obtained when g=.2, in which case the estimate is .089. However, increasing B to 200, the estimate drops to .061. (Leaving B=100 and increasing n to 30 and 40, the estimates were .072 and .06, respectively.) Thus, to be safe, B=200 or larger is recommended.

Conclusion

The main result is that for the bootstrap method studied here, among all situations considered, the estimated level of the test did not exceed .075 when testing at the .05 level provided  $B \geq 200$  is used, even with n=20. With B=100, exceptions occur, as indicated in Table 3, but given the speed of modern computers, using B=200 seems practical. In contrast, the drop-in-dispersion method requires a sample size of at

least n=100 to avoid an estimated Type I error probability greater than .075.

It was mentioned that the bootstrap method stemming from Liu and Singh (1997) was unsatisfactory in simulations; the actual probability of a Type I error was well below the nominal level. Perhaps an adjusted p-value, similar to one used here, would correct this problem in a satisfactory manner, but this has not been investigated.

Finally, R and S-Plus software is available from the author for applying the bootstrap method studied here. Ask for the function rqtest.

References

Anderson, T. W. (1958). *An introduction to multivariate statistical analysis*. New York: Wiley.

Angrist, J., Chernozhukov, V, & Fernandez-Val, I. (2006). Quantile regression under misspecification, with an application to the U.S. wage structure. *Econometrica*, 74, 539-563.

Birkes, D. & Dodge, Y. (1993). *Alternative methods of regression*. New York: Wiley.

Bradley, J. V. (1978) Robustness? *British Journal of Mathematical and Statistical Psychology*, 31, 144-152.

Buchinsky, M. (1991). *The theory and practice of quantile regression*. PhD Thesis, Dept. of Economics, Harvard University.

Hahn, J. (1995). Bootstrapping quantile regression estimators. *Econometric Theory*, 11, 105-121.

Hoaglin, D. C. (1985) Summarizing shape numerically: The g-and-h distributions. In D. Hoaglin, F. Mosteller & J. Tukey (Eds.) *Exploring data tables, trends, and shapes.*, p. 461-515. New York: Wiley.

Koenker, R. (1994). Confidence intervals for regression quantiles. In (P. Mandel M. Huskova, eds.) *Asymptotic Statistics*. Proceedings of the Fifth Prague Symposium, 349-359.

Koenker, R. & Bassett, G. (1978). Regression quantiles. *Econometrika*, 46, 33-50.

Koenker, R. & Machado, A. F. (1999). Goodness of fit related inference for quantile regression. *Journal of the American Statistical Association*, 94, 1296-1310.

Koenker, R. & Xiao, Z. J. (2002). Inference on the quantile regression process. *Econometrica*, 70, 1583-1612.

Liu, R. G. & Singh, K. (1997). Notions of limiting P values based on data depth and bootstrap. *Journal of the American Statistical Association*, 92, 266-277.

Robey, R. R. & Barcikowski, R. S. (1992). Type I error and the number of iterations in Monte Carlo studies of robustness. *British Journal of Mathematical and Statistical Psychology*, 45, 283-288.

Williams, N., Stanchina, J., Bezdjian, S., Skrok, E., Raine A. & Baker, L. (2005). *Porteus' mazes and executive function in children: Standardize Administration and scoring, and relationships to childhood aggression and delinquency*. Unpublished manuscript, Dept. of Psychology, University of Southern California.

Wilcox, R. R. (2005). *Introduction to Robust Estimation and Hypothesis Testing*, 2nd Ed. San Diego CA: Academic Press.