



On Small Sample Properties of Permutation Tests: A Significance Test for Regression Models

Tanizaki, Hisashi

(Citation)

Kobe University Economic Review, 52:27-40

(Issue Date)

2006

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/00518345>

(URL)

<https://hdl.handle.net/20.500.14094/00518345>



ON SMALL SAMPLE PROPERTIES OF PERMUTATION TESTS: A SIGNIFICANCE TEST FOR REGRESSION MODELS*

By HISASHI TANIZAKI

In this paper, we consider a nonparametric permutation test on the correlation coefficient, which is applied to a significance test on regression coefficients. Because the permutation test is very computer-intensive, there are few studies on small-sample properties, although we have numerous studies on asymptotic properties with regard to various aspects. In this paper, we aim to compare the permutation test with the t test through Monte Carlo experiments, where an independence test between two samples and a significance test for regression models are taken. For both the independence and significance tests, we obtain the results through Monte Carlo experiments that the nonparametric test performs better than the t test when the underlying sample is not Gaussian and that the nonparametric test is as good as the t test even under a Gaussian population.

1. Introduction

In the regression models, we assume that the disturbance terms are mutually independently and identically distributed. In addition, in the case where we perform the significance test on the regression coefficients, we assume that the error terms are normally distributed. Under these assumptions, it is known that the ordinary least squares (OLS) estimator of the regression coefficients follows the t distribution with $n-k$ degrees of freedom, where n and k denote the sample size and the number of regression coefficients.

As the sample size n increases, the t distribution approaches the standard normal distribution $N(0,1)$. From the central limit theorem, it is known that the OLS estimator of the regression coefficient is normally distributed for a sufficiently large sample size if the variance of the OLS estimator is finite. However, in the case where the error term is non-Gaussian and the sample size is small, the OLS estimator does not have the t distribution and therefore we cannot apply the t test. To improve these problems, in this paper we consider a significance test of the regression coefficient that includes the case where the error term is non-Gaussian and the sample size is small. A nonparametric test (or a distribution-free test) is discussed.

Generally we can regard the OLS estimator of the regression coefficient as the correlation between two samples. The nonparametric tests based on Spearman's rank correlation coefficient and Kendall's rank correlation coefficient are very famous. See, for example, Hollander and Wolfe (1973), Randles and Wolfe (1979), Conover (1980), Sprent (1989), Gibbons and Chakraborti (1992) and Hogg and Craig (1995) for the rank correlation tests. In this paper, the permutation test proposed by Fisher (1966) is utilized, and we compute the correlation coefficient for each of all the possible combinations. All the possible correlation

*This research was partially supported by Japan Society for the Promotion of Science, Grants-in-Aid for Scientific Research (C) #18530158, 2006–2009.

coefficients are compared with the correlation coefficient based on the original data. This permutation test can be directly applied to the regression problem.

The outline of this paper is as follows. In Section 2, we introduce a nonparametric test based on the permutation test, where we consider testing whether X is correlated with Y for the sample size n . Moreover, we show that we can directly apply the correlation test to the regression problem without any modification. In Section 3, we compare the powers of the nonparametric tests and the conventional t test when the underlying data are non-Gaussian. In the case where $k=2,3$ is taken for the number of regression coefficients, we examine whether the empirical sizes are correctly estimated when the significance level is $\alpha=0.10, 0.05, 0.01$.

2. The Nonparametric Test on Regression Coefficients

2.1 On Testing the Correlation Coefficient

Let $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ be a random sample, where the sample size is n . Consider testing if there is a correlation between X and Y , i.e., if the correlation coefficient ρ is zero or not. The correlation coefficient ρ is defined as:

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}},$$

where $\text{Cov}(X, Y)$, $V(X)$ and $V(Y)$ represent the covariance between X and Y , the variance of X and the variance of Y , respectively. Then, the sample correlation coefficient $\hat{\rho}$ is written as:

$$\hat{\rho} = \frac{S_{XY}}{S_X S_Y},$$

where S_{XY} , S_X and S_Y denote the sample covariance between X and Y , the sample variance of X and the sample variance of Y , which are given by:

$$S_{XY} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}), \quad S_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2, \quad S_Y^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2,$$

\bar{X} and \bar{Y} represent the sample means of X and Y .

If X is independent of Y , we have $\rho = 0$ and the joint density of X and Y is represented as a product of the marginal densities of X and Y , i.e.,

$$f_{xy}(x, y) = f_x(x) f_y(y),$$

where $f_{xy}(x, y)$, $f_x(x)$ and $f_y(y)$ denote the joint density of X and Y , the marginal density of X and the marginal density of Y . The equation above implies that for all i and j we consider randomly taking n pairs of X_i and Y_j . Accordingly, for fixed X_1 , the possible combinations are given by (X_1, Y_j) , $j=1, 2, \dots, n$, where we have n combinations. Similarly, for fixed X_2 , the possible combinations are (X_2, Y_j) , $j=2, 3, \dots, n$, i.e., $n-1$ combinations.

Moreover, we have $n-2$ combinations for X_3 , $n-3$ combinations for X_4 and so on. Therefore, the total number of possible combinations between X and Y are given by $n!$. For each combination, we can compute the correlation coefficient. Thus, $n!$ correlation coefficients are obtained. The $n!$ correlation coefficients are compared with the correlation coefficient obtained from the original pairs of data. If the correlation coefficient obtained from the original data is in the tail of the empirical distribution constructed from the $n!$ correlation coefficients, the hypothesis that X is correlated with Y is rejected. The testing procedure above is distribution-free or nonparametric, and can be applied in almost all cases. The nonparametric test discussed above is known as a permutation test, which has developed by Fisher (1966). For example, see Stuart and Ord (1991).

The order of X_i , $i=1,2,\dots, n$, is fixed and we permute Y_j , $j=1,2,\dots, n$, randomly. Based on the $n!$ correlation coefficients, we can test if X is correlated with Y . Let the $n!$ correlation coefficients be $\hat{\rho}^{(i)}$, $i=1,2,\dots, n!$. Suppose that $\hat{\rho}^{(1)}$ is the correlation coefficient obtained from the original data. The estimator of the correlation coefficient ρ , denoted by $\hat{\rho}$, is distributed as:

$$P(\hat{\rho} < \hat{\rho}^{(1)}) = \frac{\text{Number of combinations less than } \hat{\rho}^{(1)} \text{ out of } \hat{\rho}^{(1)}, \dots, \hat{\rho}^{(n!)}}{\text{Number of all possible combinations (i.e., } n!)},$$

$$P(\hat{\rho} = \hat{\rho}^{(1)}) = \frac{\text{Number of combinations equal to } \hat{\rho}^{(1)} \text{ out of } \hat{\rho}^{(1)}, \dots, \hat{\rho}^{(n!)}}{\text{Number of all possible combinations (i.e., } n!)},$$

$$P(\hat{\rho} > \hat{\rho}^{(1)}) = \frac{\text{Number of combinations greater than } \hat{\rho}^{(1)} \text{ out of } \hat{\rho}^{(1)}, \dots, \hat{\rho}^{(n!)}}{\text{Number of all possible combinations (i.e., } n!)}.$$

Thus, the above three probabilities can be computed. The null hypothesis $H_0: \rho = 0$ is rejected by the two-sided test if $P(\hat{\rho} < \hat{\rho}^{(1)})$ or $P(\hat{\rho} > \hat{\rho}^{(1)})$ is small enough.

Note as follows. S_{XY} is rewritten as:

$$S_{XY} = \frac{1}{n} \sum_{i=1}^n X_i Y_i - \bar{X}\bar{Y}.$$

The sample means \bar{X} and \bar{Y} take the same values without depending on the order of X and Y . Similarly, S_X and S_Y are independent of the order of X and Y . Therefore, $\hat{\rho}$ depends on $\sum_{i=1}^n X_i Y_i$. That is, for the empirical distribution based on the $n!$ correlation coefficients, $\hat{\rho}$ is a monotone function of $\sum_{i=1}^n X_i Y_i$, which implies that we have a one-to-one correspondence between $\hat{\rho}$ and $\sum_{i=1}^n X_i Y_i$. Therefore, for $\sum_{i=1}^n X_i Y_i$, we may compute the $n!$ combinations by changing the order of Y_i , $i=1,2,\dots, n$. Thus, by utilizing $\sum_{i=1}^n X_i Y_i$ rather than $\hat{\rho}$, the computational burden can be reduced.

As for a special case, suppose that (X_i, Y_i) , $i=1,2,\dots, n$, are normally distributed, i.e.,

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix} \right).$$

Under the null hypothesis $H_0: \rho = 0$, the sample correlation coefficient $\hat{\rho}$ is distributed as the following t distribution:

$$\frac{\hat{\rho}\sqrt{n-2}}{\sqrt{1-\hat{\rho}^2}} \sim t(n-2),$$

Note that we cannot use a t distribution in the case of testing the null hypothesis $H_0: \rho = \rho_0$. For example, see Lehmann (1986), Stuart and Ord (1991, 1994) and Hogg and Craig (1995). Generally, it is natural to consider that (X, Y) is non-Gaussian and that the distribution of (X, Y) is not known. If the underlying distribution is not Gaussian but the t distribution is applied to the null hypothesis $H_0: \rho = 0$, the appropriate testing results cannot be obtained. However, the nonparametric permutation test can be applied even in the Non-Gaussian cases, because it is distribution-free.

2.2. On Testing the Regression Coefficient

Using exactly the same approach as the nonparametric test on the correlation coefficient, discussed in Section 2.1, we consider a nonparametric significance test on the regression coefficients.

The regression model is given by:

$$Y_i = X_i \beta + u_i, \quad i=1, 2, \dots, n,$$

where the OLS estimator of β , i.e., $\hat{\beta}$, is represented as:

$$\hat{\beta} = (X'X)^{-1} X'Y = \sum_{i=1}^n (X'X)^{-1} X'_i Y_i. \quad (1)$$

Note as follows:

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, \quad X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}, \quad \hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_n \end{pmatrix},$$

where Y_i denotes the i -th element of a $n \times 1$ vector Y and X_i indicates the i -th row vector of a $n \times k$ matrix X .

From the structure of equation (1), when X_i in Section 2.1 is replaced by $(X'X)^{-1}X'_i$, we can find that the same discussion as in Section 2.1 holds with only one difference, i.e., X_i is a scalar in Section 2.1 while $(X'X)^{-1}X'_i$ is a $k \times 1$ vector in this section. We have $n!$ regression coefficients by changing the order of Y . Let $\hat{\beta}^{(i)}$, $i=1, 2, \dots, n!$, be the $n!$ regression coefficients and $\hat{\beta}_j^{(i)}$ be the j -th element of $\hat{\beta}^{(i)}$, i.e., $\hat{\beta}^{(i)} = (\hat{\beta}_1^{(i)}, \hat{\beta}_2^{(i)}, \dots, \hat{\beta}_K^{(i)})$. Suppose that $\hat{\beta}_j^{(1)}$ represents the j -th element of the regression coefficient vector obtained from the original data series. Under the null hypothesis $H_0: \beta_j = 0$, the empirical distribution of $\hat{\beta}_j$, which is the j -th element of the OLS

estimator of β , is given by:

$$\begin{aligned} P(\hat{\beta}_j < \hat{\beta}_j^{(1)}) &= \frac{\text{Number of combinations less than } \hat{\beta}_j^{(1)} \text{ out of } \hat{\beta}_j^{(1)}, \dots, \hat{\beta}_j^{(n!)}}{\text{Number of all possible combinations (i.e., } n!)}, \\ P(\hat{\beta}_j = \hat{\beta}_j^{(1)}) &= \frac{\text{Number of combinations equal to } \hat{\beta}_j^{(1)} \text{ out of } \hat{\beta}_j^{(1)}, \dots, \hat{\beta}_j^{(n!)}}{\text{Number of all possible combinations (i.e., } n!)}, \\ P(\hat{\beta}_j > \hat{\beta}_j^{(1)}) &= \frac{\text{Number of combinations greater than } \hat{\beta}_j^{(1)} \text{ out of } \hat{\beta}_j^{(1)}, \dots, \hat{\beta}_j^{(n!)}}{\text{Number of all possible combinations (i.e., } n!)}. \end{aligned}$$

For all $j=1,2,\dots,k$, we can implement the same computational procedure as above and compute each probability. We can perform the significance test by examining where $\hat{\beta}_j^{(1)}$ is located among the $n!$ regression coefficients. The null hypothesis $H_0: \beta_j = 0$ is rejected by the two-sided test if $P(\beta_j < \hat{\beta}_j^{(1)})$ or $P(\beta_j > \hat{\beta}_j^{(1)})$ is small enough.

Generally, as for the testing procedure of the null hypothesis $H_0: \beta = \beta^*$, we may consider a nonparametric permutation test between $(X'X)^{-1} X_i'$ and $Y_i - X_i \beta^*$, because $\hat{\beta} - \beta$ is transformed into:

$$\begin{aligned} \hat{\beta} - \beta &= (X'X)^{-1} X_i' Y - \beta \\ &= (X'X)^{-1} X' (Y - X\beta) \\ &= \sum_{i=1}^n (X'X)^{-1} X_i' (Y_i - X_i \beta). \end{aligned}$$

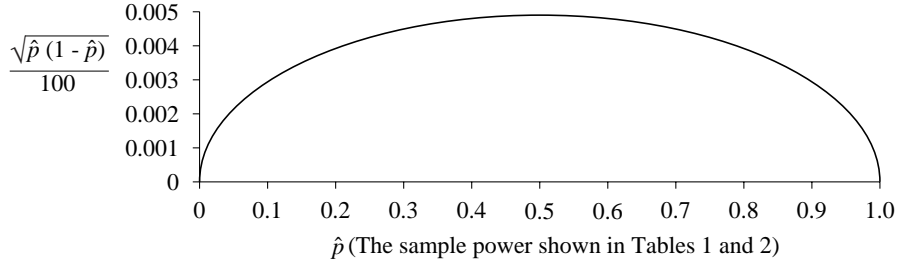
Note as follows. As for the conventional parametric significance test, the error terms u_i , $i = 1, 2, \dots, n$, are assumed to be mutually independently and normally distributed with mean zero and variance σ^2 . Under the null hypothesis $H_0: \beta_j = \beta_j^*$, the j -th element of the OLS estimator (i.e., $\hat{\beta}_j$) is distributed as:

$$\frac{\hat{\beta}_j - \beta_j^*}{S\sqrt{a_{jj}}} \sim t(n-k),$$

where a_{jj} denotes the j -th diagonal element of $(X'X)^{-1}$. β^* and S^2 represent $\beta^* = (\beta_1^*, \beta_2^*, \dots, \beta_k^*)'$ and $S^2 = (Y - X\hat{\beta})'(Y - X\hat{\beta})/(n-k)$, respectively. Thus, only when u_i is assumed to be normal, we can use the t distribution. However, unless u_i is normal, the conventional t test gives us the incorrect inference in the small sample. As it is well known, in a large sample $\sqrt{n}(\hat{\beta}_j - \beta_j)$ is asymptotically normal when the variance of u_i is finite. Thus, the case of a large sample is different from that of a small sample. In this paper, under the non-Gaussian assumption, we examine the powers of the nonparametric tests on the correlation coefficient through Monte Carlo experiments. Moreover, in the regression analysis we examine how robust the conventional t test is when the underlying population is not Gaussian.

TABLE 1. Empirical Sizes and Sample Powers ($H_0 : \rho = 0$ and $H_1 : \rho \neq 0$)

n	ρ	α	Nonparametric Permutation Test					Parametric t Test				
			N	X	U	L	C	N	X	U	L	C
6	.0	.10	.1008	.0986	.1040	.1035	.1030	.0981	.1019	.1019	.0987	.1180
		.05	.0499	.0509	.0512	.0518	.0515	.0474	.0646	.0507	.0483	.0748
		.01	.0112	.0096	.0113	.0101	.0098	.0099	.0225	.0115	.0093	.0269
	.3	.10	.1491	.1830	.1479	.1621	.2605	.1496	.2284	.1488	.1585	.3055
		.05	.0821	.1108	.0798	.0926	.1639	.0795	.1638	.0794	.0914	.2375
		.01	.0189	.0323	.0191	.0205	.0475	.0203	.0692	.0188	.0225	.1389
	.6	.10	.2876	.3199	.2605	.3014	.3834	.2932	.3814	.2646	.3081	.4384
		.05	.1776	.2145	.1528	.1836	.2696	.1787	.2917	.1568	.1939	.3633
		.01	.0460	.0755	.0386	.0501	.0983	.0481	.1579	.0389	.0564	.2405
	.9	.10	.4674	.4469	.4365	.4682	.4802	.4792	.5144	.4475	.4810	.5325
		.05	.3117	.3206	.2786	.3240	.3636	.3300	.4208	.2887	.3410	.4549
		.01	.0894	.1269	.0745	.1009	.1572	.1060	.2501	.0843	.1240	.3096
8	.0	.10	.1018	.0994	.1040	.1014	.1007	.1020	.0986	.1057	.1025	.1183
		.05	.0523	.0489	.0523	.0549	.0494	.0515	.0609	.0545	.0549	.0791
		.01	.0103	.0111	.0110	.0100	.0093	.0100	.0225	.0126	.0113	.0345
	.3	.10	.1869	.2025	.1748	.1971	.3120	.1866	.2514	.1779	.1978	.3471
		.05	.1066	.1236	.0940	.1096	.2107	.1083	.1821	.0998	.1147	.2780
		.01	.0257	.0385	.0211	.0251	.0749	.0248	.0891	.0250	.0254	.1814
	.6	.10	.3932	.3997	.3605	.4002	.4527	.3996	.4592	.3657	.4049	.4963
		.05	.2602	.2833	.2212	.2611	.3455	.2693	.3695	.2359	.2742	.4205
		.01	.0815	.1148	.0627	.0850	.1615	.0865	.2149	.0714	.0959	.2920
	.9	.10	.6197	.5445	.6256	.6209	.5647	.6289	.6023	.6309	.6313	.5979
		.05	.4653	.4163	.4511	.4850	.4684	.4806	.5125	.4664	.4980	.5315
		.01	.1966	.2034	.1562	.2119	.2646	.2199	.3422	.1795	.2395	.4096
10	.0	.10	.1046	.0987	.0999	.0984	.0992	.1033	.0943	.1006	.0950	.1058
		.05	.0524	.0486	.0520	.0489	.0477	.0510	.0570	.0526	.0479	.0733
		.01	.0106	.0083	.0104	.0102	.0090	.0097	.0220	.0114	.0098	.0348
	.3	.10	.2162	.2243	.2046	.2228	.3468	.2151	.2765	.2065	.2231	.3764
		.05	.1277	.1429	.1179	.1348	.2513	.1284	.2046	.1218	.1348	.3133
		.01	.0338	.0484	.0285	.0351	.1022	.0355	.1037	.0306	.0367	.2149
	.6	.10	.4766	.4501	.4640	.4920	.5089	.4775	.5005	.4692	.4954	.5378
		.05	.3387	.3335	.3118	.3579	.4081	.3416	.4149	.3198	.3599	.4670
		.01	.1325	.1522	.1020	.1424	.2281	.1375	.2604	.1106	.1504	.3610
	.9	.10	.7336	.6128	.7539	.7276	.6075	.7358	.6663	.7558	.7348	.6343
		.05	.6008	.4958	.6038	.6032	.5258	.6116	.5806	.6106	.6129	.5734
		.01	.3116	.2812	.2719	.3286	.3318	.3295	.4128	.2925	.3482	.4553

FIGURE 1. Standard Error of \hat{p}

3. Monte Carlo Experiments

3.1 On Testing the Correlation Coefficient

Each value in Table 1 represents the rejection rates of the null hypothesis $H_0: \rho = 0$ against the alternative hypothesis $H_1: \rho \neq 0$ (i.e., the two-sided test is chosen) by the significance level $\alpha = 0.01, 0.05, 0.10$, where the experiment is repeated 10^4 times. That is, in Table 1 the number which the correlation coefficient obtained from the original observed data is less than $\alpha/2$ or greater than $1 - \alpha/2$ is divided by 10^4 . In other words, we compute the probabilities $P(\hat{\rho} < \hat{\rho}^{(1)})$ and $P(\hat{\rho} > \hat{\rho}^{(1)})$ and repeat the experiment 10^4 times. The probabilities correspond to the sample powers, where $\hat{\rho}^{(1)}$ denotes the correlation coefficient computed from the original data. The ratio of $P(\hat{\rho} < \hat{\rho}^{(1)}) \leq \alpha/2$ or $P(\hat{\rho} > \hat{\rho}^{(1)}) \leq \alpha/2$ is shown in Table 1. $\alpha = 0.10, 0.05, 0.01$ is examined. N, X, U, L and C indicate the distributions of (X, Y) , which denote the standard normal distribution $N(0,1)$, the chi-squared distribution $\chi^2(1)-1$, the uniform distribution $U(-1,1)$, the logistic distribution $e^{-x}/(1+e^{-1})^2$, and the Cauchy distribution $(\pi(1+x^2))^{-1}$, respectively. The standard error of the empirical power, denoted by \hat{p} , is obtained by $\sqrt{\hat{p}(1-\hat{p})}/100$, which is displayed in Figure 1. For example, when $\hat{p} = 0.5$ the standard error takes the maximum value, which is 0.005.

In this paper, the random draws of (X_i, Y_i) are obtained as follows. Let u_i and v_i be the random variables which are mutually independently distributed. Each of u_i and v_i is generated as the standard normal random variable $N(0,1)$, the chi-squared random variable $\chi^2(1)-1$, the uniform random variable $U(-1,1)$, the logistic distribution $e^{-x}/(1+e^{-1})^2$, or the Cauchy distribution $(\pi(1+x^2))^{-1}$. Denote the correlation coefficient between X and Y by ρ . Given the random draws of (u_i, v_i) and ρ , (X_i, Y_i) is transformed into:

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} = \begin{pmatrix} 1 & \rho \\ 0 & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} u_i \\ v_i \end{pmatrix}.$$

In the case of the Cauchy distribution the correlation coefficient does not exist because the Cauchy random variable has neither mean nor variance. Even in the case of the Cauchy distribution, however, we can obtain the random draws of (X_i, Y_i) given (u_i, v_i) and ρ , utilizing the above formula.

Using the artificially generated data given the true correlation coefficient $\rho = 0.0, 0.3, 0.6, 0.9$, we test the null hypothesis $H_0: \rho = 0$ against the alternative hypothesis $H_1: \rho \neq 0$. Taking the significance level $\alpha = 0.10, 0.05, 0.01$ and the two-sided test, the rejection rates out of 10^4 experiments are shown in Table 1. In addition, taking the sample size $n = 6, 8, 10$, both the nonparametric permutation test and the parametric t test are reported in the table.

As it is easily expected, for the normal sample (i.e., N), the t test performs better than the nonparametric test, but for the other samples (i.e., X, U, L and C), the nonparametric test is superior to the t test.

Each value in the case of $\rho = 0$ represents the empirical size, which should be theoretically equal to the significance level α . The results are as follows. In the case of N and L, each value is quite close to α for the t test. However, for all $n = 6, 8, 10$, the t tests of X and C are over-rejected especially in the case of $\alpha = 0.05, 0.01$. Taking an example of $n = 6$, X is 0.0646 and C is 0.0748 when $\alpha = 0.05$, while X is 0.0225 and C is 0.0269 when $\alpha = 0.01$. The t test of U is also slightly over-rejected in the case of $\alpha = 0.05, 0.01$. Thus, we often have the case where the t test over-rejects the null hypothesis $H_0: \rho = 0$ depending on the underlying distribution. However, the nonparametric test rejects the null hypothesis with probability α for all the underlying distributions. Accordingly, through a comparison of the empirical size, we can conclude that the nonparametric test is more robust than the t test.

Next, we examine the cases of $\rho = 0.3, 0.6, 0.9$ to compare the sample powers of the two tests (note that each value of $\rho = 0.3, 0.6, 0.9$ in Table 1 corresponds to the sample power). As for X and C, it is not meaningful to compare the sample powers because the empirical sizes are already over-estimated. Regarding the sample powers of N, U and L, the nonparametric test is close to the t test. Especially, for N, it is expected that the t test is better than the nonparametric test, but we can see that the nonparametric test is as good as the t test in the sense of the sample power.

Thus, the permutation-based nonparametric test introduced in this paper is useful because it gives us the correct empirical size and is powerful even though it does not need to assume the distribution function.

3.2 On Testing the Regression Coefficient

In Table 2, the testing procedure taken in Section 3.1 is applied to the regression analysis. Let $X_i = (X_{1,i}, X_{2,i}, \dots, X_{k,i})$, where $X_{1,i} = 1$ and $X_{j,i} \sim N(0,1)$ for $j = 2, 3$ are set. $X_{2,i}$ and $X_{3,i}$ are mutually independently generated. The error term u_i is assumed to have the standard normal distribution $N(0,1)$, the chi-squared distribution $\chi^2(1)$, the uniform distribution $U(-1,1)$, the logistic distribution $e^{-x}/(1+e^{-1})^2$, or the Cauchy distribution $(\pi(1+x^2))^{-1}$. Under the null hypothesis $H_0: \beta = 0$, we obtain $y_i = u_i$. Therefore, we consider the correlation between $(X'X)^{-1}X_i$ and y_i .

The sample size is $n = 6, 8, 10$ and the number of the regression coefficient to be estimated is $k = 2, 3$. The nonparametric test is compared with the t test in the criterion of the empirical size. Each value in Table 2 represents the rejection rate out of 10^4 simulation runs. Therefore,

theoretically each value in the table should be equivalent to the probability which rejects the null hypothesis $H_0 : \beta_j = 0$ against the alternative hypothesis $H_1 : \beta_j \neq 0$ for $j=1,2,3$, which probability corresponds to the significance level α .

As in Section 3.1, in the case of N, the t test should be better than the nonparametric test, i.e., the t test should be closer to the significance level α than the nonparametric test, because the OLS estimator of β_j follows the t distribution with $n - k$ degrees of freedom when the error terms u_i , $i=1,2,\dots, n$, are mutually independently and normally distributed with mean zero.

The superscripts * and ** in the table imply that the empirical size is statistically different from α by the two-sided test of the significance levels 5% and 1%, respectively. That is, each value without the superscript * and ** gives us the correct size. The superscripts * or ** are put as follows. By the central limit theorem, we can approximate as $(\hat{p} - \alpha)/\sqrt{p(1-p)/G} \sim N(0,1)$ when G is large enough, where p is a parameter and \hat{p} denotes its estimate. G represents the number of simulation runs. $G=10^4$ is taken in this paper. Under the null hypothesis $H_0 : p = \alpha$ against the alternative hypothesis $H_1 : p \neq \alpha$, the test statistic $(\hat{p} - \alpha)/\sqrt{p(1-\alpha)/G}$ is asymptotically normally distributed. Thus, we can test whether p is different from α .

The results are as follows. For the t tests of X and U, the constant term (i.e., β_1) gives us the over-estimated empirical sizes in both cases of $k = 2,3$. On the contrary, for the t tests of C, the constant term (i.e., β_1) yields the under-estimated empirical sizes. However, for the nonparametric tests, almost all the values in Table 2 are very close to the significance level α . Therefore, we can conclude that the nonparametric test is superior to the t test in the sense of the corrected empirical size.

Power Comparison: Next, we compare the sample powers in the case of $k = 3$. Figures 2-4 display differences of the sample powers between the two tests for $\beta_3 = 0.1, 0.2, \dots, 0.9$, where $\beta_1 = \beta_2 = 0$ and $\alpha = 0.1$ are taken. Since the sample powers of the permutation test are subtracted from those of the t test, the values above the horizontal line imply that the permutation test is more powerful than the t test. Take the case $n = 6$ as an example, i.e., Figure 2. For all β_3 of L and C, the differences are positive and accordingly we can conclude that for L and C the permutation test is superior to the t test. Furthermore, the differences are positive for $\beta_3 = 0.1 - 0.5$ of N and X but negative for β_3 greater than 0.6. Thus, we can observe through Figures 2-4 that except for a few cases the permutation test is more powerful than the t test for all $n = 6, 8, 10$ at least when β_3 is small. From the three figures, in addition, each difference becomes small as n is large. Therefore, the permutation test is close to the t test as the sample size is large.

CPU Time: As mentioned above, in the case where we perform the significance test of the regression coefficient, we need to compute the $n!$ regression coefficients (for example, $n!$ is equal to about 3.6 million when $n = 10$). In Table 3, CPU time in one simulation run is shown for $n = 11, 12, 13, 14$ and $k = 2, 3, 4$, where a Pentium III 1GHz CPU personal computer and a WATCOM Fortran 77/32 Compiler (Version 11.0) are utilized. The order of computation is

TABLE 2. Significance Tests on Regression Coefficients: Empirical Sizes

	n	α	$k = 2$				$k = 3$					
			Nonparametric Test		t Test		Nonparametric Test			t Test		
			β_1	β_2	β_1	β_2	β_1	β_2	β_3	β_1	β_2	β_3
N	6	.10	.1008	.1008	.1055	.0981	.1076**	.1075**	.1046	.1041	.1017	.1033
		.05	.0500	.0499	.0534	.0474	.0575	.0546**	.0543**	.0520	.0492	.0512
		.01	.0113	.0112	.0104	.0099	.0105	.0110	.0111	.0105	.0098	.0095
	8	.10	.0964	.0963	.0989	.0967	.1024	.1016	.0991	.0989	.1000	.0971
		.05	.0443**	.0443**	.0454**	.0454**	.0516	.0503	.0505	.0473	.0489	.0503
		.01	.0087	.0087	.0093	.0080**	.0106	.0079	.0101	.0098	.0088	.0112
	10	.10	.0988	.0985	.0983	.0988	.0989	.0976	.1023	.0997	.0979	.1029
		.05	.0503	.0502	.0487	.0493	.0489	.0488	.0508	.0490	.0510	.0524
		.01	.0115	.0114	.0103	.0112	.0112	.0102	.0105	.0122**	.0098	.0122**
X	6	.10	.1013	.1013	.2132**	.0991	.1061**	.1042	.1073**	.1964**	.1011	.0999
		.05	.0517	.0516	.1569	.0479	.0546**	.0533	.0546**	.1318**	.0508	.0495
		.01	.0109	.0109	.0806	.0099	.0125**	.0121**	.0110	.0506**	.0102	.0108
	8	.10	.1029	.1028	.1967**	.1034	.0960	.1038	.1013	.1867**	.1036	.0996
		.05	.0521	.0521	.1468**	.0520	.0469	.0530	.0504	.1282**	.0533	.0507
		.01	.0108	.0108	.0794**	.0116	.0092	.0101	.0101	.0586**	.0116	.0111
	10	.10	.1029	.1028	.1787**	.1027	.1028	.1059**	.0998	.1753**	.1048	.1031
		.05	.0493	.0491	.1321**	.0492	.0502	.0530	.0490	.1215**	.0560**	.0482
		.01	.0095	.0094	.0729**	.0092	.0104	.0099	.0102	.0611**	.0097	.0109
U	6	.10	.1082**	.1082**	.1125**	.1042	.1003	.0992	.1042	.1038	.0996	.1039
		.05	.0565**	.0564**	.0621**	.0541	.0515	.0516	.0551**	.0554**	.0521	.0527
		.01	.0112	.0112	.0155**	.0106	.0118	.0113	.0116	.0121**	.0111	.0087
	8	.10	.0981	.0981	.0962	.0997	.0947	.0988	.0968	.1067**	.1026	.1008
		.05	.0490	.0490	.0543**	.0480	.0477	.0488	.0469	.0550**	.0523	.0490
		.01	.0108	.0108	.0145**	.0100	.0088	.0088	.0093	.0129**	.0102	.0105
	10	.10	.0953	.0953	.0980	.0958	.0983	.0980	.1027	.0988	.0994	.1015
		.05	.0493	.0493	.0525	.0487	.0474	.0493	.0499	.0528	.0496	.0495
		.01	.0102	.0102	.0131**	.0094	.0082	.0098	.0098	.0125**	.0091	.0099
L	6	.10	.1080**	.1080**	.0997	.1054	.1012	.1004	.1006	.0985	.0989	.1010
		.05	.0562**	.0562**	.0491	.0524	.0503	.0507	.0535	.0476	.0516	.0517
		.01	.0118	.0118	.0080**	.0095	.0125**	.0119	.0124**	.0081	.0111	.0094
	8	.10	.0984	.0984	.0906**	.1005	.0964	.1024	.0979	.1018	.1052	.1007
		.05	.0505	.0505	.0448**	.0508	.0487	.0504	.0475	.0476	.0515	.0496
		.01	.0111	.0111	.0084	.0106	.0108	.0094	.0091	.0096	.0101	.0111
	10	.10	.0933**	.0933**	.0946	.0934**	.0989	.1007	.1018	.0991	.0993	.1011
		.05	.0492	.0492	.0463	.0474	.0495	.0506	.0517	.0467	.0510	.0509
		.01	.0110	.0110	.0082	.0107	.0093	.0095	.0106	.0088	.0095	.0101
C	6	.10	.1114**	.1114**	.0709**	.1035	.0967	.1004	.1021	.0847**	.0984	.1028
		.05	.0566**	.0566**	.0303**	.0505	.0510	.0520	.0505	.0375**	.0501	.0517
		.01	.0116	.0116	.0042**	.0104	.0122**	.0116	.0119	.0066**	.0094	.0098
	8	.10	.0983	.0983	.0642**	.1013	.0998	.1049	.0975	.0802**	.1087**	.1029
		.05	.0486	.0486	.0259**	.0499	.0500	.0514	.0475	.0361**	.0530	.0507
		.01	.0103	.0103	.0054**	.0116	.0110	.0097	.0092	.0065**	.0106	.0102
	10	.10	.0958	.0957	.0693**	.0957	.0976	.1016	.1040	.0803**	.1044	.1032
		.05	.0461	.0461	.0261**	.0484	.0488	.0503	.0513	.0327**	.0512	.0516
		.01	.0097	.0097	.0023**	.0101	.0097	.0097	.0108	.0051**	.0095	.0097

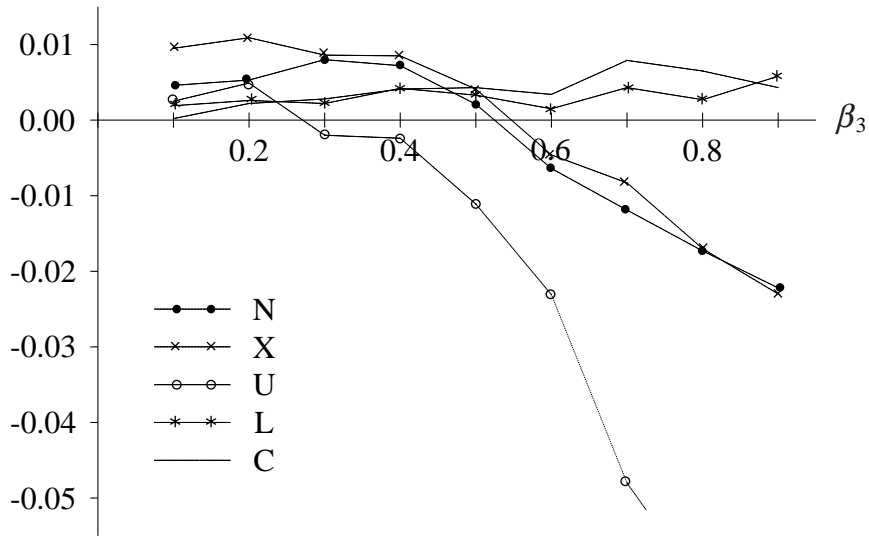


FIGURE 2. Difference of the Sample Powers between the Two Tests
($n = 6$, $k = 3$, $\beta_1 = \beta_2 = 0$ and $\alpha = 0.1$)

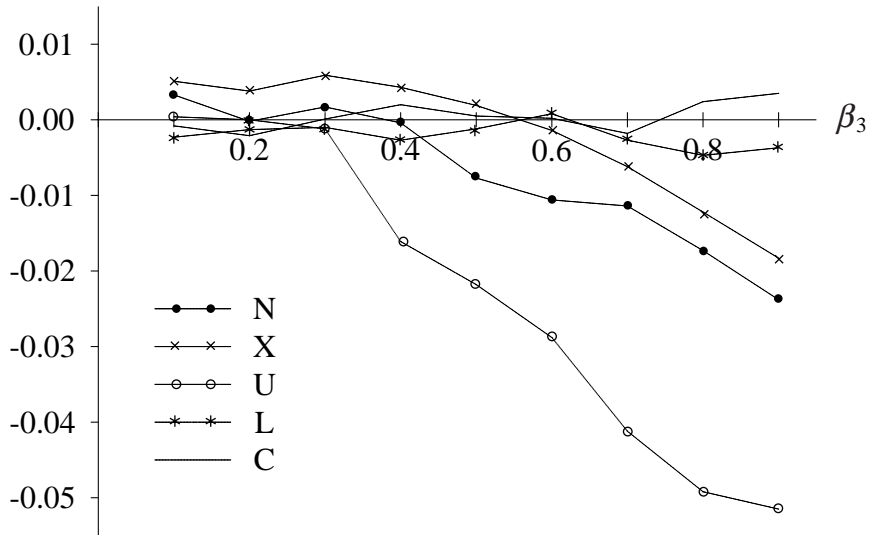


FIGURE 3. Difference of the Sample Powers between the Two Tests
($n = 8$, $k = 3$, $\beta_1 = \beta_2 = 0$ and $\alpha = 0.1$)

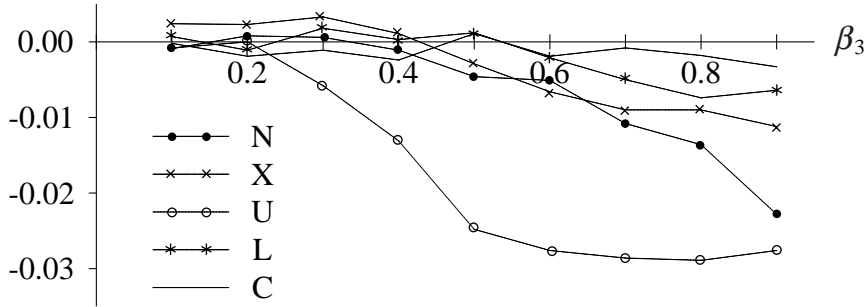


FIGURE 4. Difference of the Sample Powers between the Two Tests
($n = 10$, $k = 3$, $\beta_1 = \beta_2 = 0$ and $\alpha = 0.1$)

TABLE 3. CPU Time (minutes)

n / k	2	3	4
11	0.18	0.26	0.33
12	1.94	2.67	3.42
13	26.73	37.21	48.24
14	377.32	527.90	685.49

about $n! \times k$. The case of sample size n is n times more computer-intensive than that of sample size $n - 1$. For example, it might be expected that the case of $n = 15$ and $k = 4$ takes about one week (i.e., 15×685.49 minutes) to obtain a result. This is not feasible in practice. Thus, the permutation test discussed in this paper is very computer-intensive.

Therefore, we need to consider less computationally intensive procedure. In order to reduce the computational burden when $n!$ is large, it might be practical to choose some of the $n!$ combinations and perform the same testing procedure discussed in this paper. That is, taking N combinations out of the $n!$ combinations randomly, we compute the probabilities $P(\hat{\beta}_j < \hat{\beta}_j^{(1)})$ and $P(\hat{\beta}_j > \hat{\beta}_j^{(1)})$. If either of them is smaller than $\alpha = 0.1$, the null hypothesis $H_0: \beta_j = 0$ is rejected. We examine whether empirical sizes depend on N . The results are in Table 4, where $k = 2$ and $N = 10^4, 10^5, 10^6, 10!$ are taken. The case of $N = 10!$ corresponds to the nonparametric test of β_2 in the case of $\alpha = 0.1$ and $k = 2$ of Table 2, where the empirical sizes are computed using all the possible combinations. All the cases of $N = 10^4, 10^5, 10^6$ are very close to those of $N = 10!$. $N = 10^4$ is sufficiently large in this case, but more than $N = 10^4$ should be taken for safety. Thus, we can choose some out of the $n!$ combinations and compute the corresponding probabilities. A less computationally intensive procedure might be possible when n is large, i.e., when n is larger than 14 in my case.

TABLE 4. Empirical Sizes: $n = 10$, $k = 2$, $\beta_1 = \beta_2 = 0$ and $\alpha = 0.1$

N	10^4	10^5	10^6	$10!$
N	.0985	.1020	.0994	.0985
X	.0962	.1010	.1031	.1028
U	.0936	.0971	.0980	.0953
L	.0937	.0982	.0967	.0933
C	.0968	.1026	.0979	.0957

4. Summary

Only when the error term is normally distributed, we can utilize the t test for testing the regression coefficient. Since the distribution of the error term is not known, we need to check whether the normality assumption is plausible before testing the hypothesis. As a result of testing, in the case where the normality assumption is rejected, we cannot test the hypothesis on the regression coefficient using the t test. In order to improve this problem, in this paper we have shown a significance test on the regression coefficient, which can be applied to any distribution.

In Section 3.1, we tested whether the correlation coefficient between two samples is zero and examined the sample powers of the two tests. For each of the cases where the underlying samples are normal, chi-squared, uniform, logistic and Cauchy, 10^4 simulation runs are performed and the nonparametric permutation test is compared with the parametric t test with respect to the empirical sizes and the sample powers. As it is easily expected, the t test is sometimes a biased test under the non-Gaussian assumption. That is, we have the cases where the empirical sizes are over-estimated. However, the nonparametric test gives us the correct empirical sizes without depending on the underlying distribution. Specifically, even when the sample is normal, the nonparametric test is very close to the t test (theoretically, the t test should be better than any other test when the sample is normal).

In Section 3.2, we have performed Monte Carlo experiments on the significance test of the regression coefficients. It might be concluded that the nonparametric test is closer to the true size than the t test for almost all the cases. Moreover, the sample powers are compared for both tests. As a result, we find that the permutation test is more powerful than the t test when the null hypothesis is close to the alternative hypothesis, i.e., when β_3 is small in Figures 2 - 4. Thus, we find through the Monte Carlo experiments that the nonparametric test discussed in this paper can be applied to any distribution of the underlying sample. However, the problem is that the nonparametric test is too computer-intensive. We have shown that when $n!$ is too large it is practical to choose some of the $n!$ combinations and perform the testing procedure. Taking an example of $n = 10$ and $k = 2$, we have obtained the result that we can perform the testing procedure taking the 10^4 combinations out of all the possible combinations ($10!$ in this case) randomly. Thus, it has been shown in this paper that we can reduce the computational burden.

REFERENCES

- W.J. Conover (1980), *Practical Nonparametric Statistics* (Second Edition), John Wiley & Sons.
- Fisher, R.A. (1966) *The Design of Experiments* (eighth edition), New York: Hafner.
- J.D. Gibbons and S. Chakraborti (1992), *Nonparametric Statistical Inference* (Third Edition, revised and Expanded), Marcel Dekker.
- R.V. Hogg and A.T. Craig (1995), *Introduction to Mathematical Statistics* (Fifth Edition), Prentice Hall.
- M. Hollander and D.A. Wolfe (1973), *Nonparametric Statistical Methods*, John Wiley & Sons.
- E.L. Lehmann (1986), *Testing Statistical Hypotheses* (Second Edition), John Wiley & Sons.
- R.H. Randles and D.A. Wolfe (1979), *Introduction to the Theory of Nonparametric Statistics*, John Wiley & Sons.
- P. Sprent (1989), *Applied Nonparametric Statistical Methods*, Chapman and Hall.
- A. Stuart and J.K. Ord (1991), *Kendall's Advanced Theory of Statistics, Vol.2* (Fifth Edition), Edward Arnold.
- A. Stuart and J.K. Ord (1994), *Kendall's Advanced Theory of Statistics, Vol.1* (Sixth Edition), Edward Arnold.