American Journal of Theoretical and Applied Statistics
2025; 14(2): 61-76
http://www.sciencepublishinggroup.com/j/ajtas
doi: 10.11648/j.ajtas.20251402.12
ISSN: 2326-8999 (Print); ISSN: 2326-9006 (Online)

Study on Efficacy of Kendall's τ Based Test Statistic in Generalized Partially Linear Regression Model

Sthitadhi Das

Department of Mathematics, Brainware University, Kolkata, India

Email address:
sthdas999@gmail.com (Sthitadhi Das)

To cite this article:
Sthitadhi Das. (2025). Study on Efficacy of Kendall's τ Based Test Statistic in Generalized Partially Linear Regression Model. *American Journal of Theoretical and Applied Statistics, 14*(2), 61-76. https://doi.org/10.11648/j.ajtas.20251402.12

Received: 12 February 2025; **Accepted:** 3 March 2025; **Published:** 24 March 2025

Abstract: Under the setup of a generalized partially linear model $Y = \beta_1 X_1 + \ldots + \beta_p X_p + m(W_1, \ldots, W_q) + \epsilon$ with p parametric regressors $X_1, \ldots, X_p$ and q nonparametric regressors $W_1, \ldots, W_q$, we are motivated to test on the independence between all the $(p + q)$ regressors and the random error ϵ. Since one obtains unbiased prediction of the study variable of interest when the regressors under consideration are independent of the random error, such testing of independence is a vital objective indeed. Here, $m(W_1, \ldots, W_q)$ is a Lipschitz continuous function defined on $\mathbb{R}^q \to \mathbb{R}$. To carry out the prescribed testing scheme, some test statistics are formed based on the nonparametric measure of association Kendall's (1938) τ. Moreover, as an implication of the null hypothesis suggesting independence between joint $(p + q)$ regressors and ϵ, we further modify the hypothesis as ϵ is not observable at all. Eventually, independence among the r-th order difference of estimated response and the r-th order difference of observed response is implied from the original null hypothesis. Later, the concept of V-statistic is applied to propose the test statistics based on the paired observations on the r-th differences of the estimated and observed Y. Their consistent power performances are achieved under a sequence of contiguous alternatives stating complete dependence between error and regressors while testing the independence of regressors with ϵ. Le Cam (1960)'s theory on contiguity is required to develop a sequence of contiguous alternative hypotheses in this regard. The asymptotic powers of the test statistics are evaluated with the help of their asymptotic distributions under null hypothesis of independence and contiguous alternatives. Subsequently, a data analysis is performed to substantiate the eligibility of the proposed test statistics in such testing setup.

Keywords: Generalized Partially Linear Regression Model, Nonparametric Regression Model, Kendall's τ, Measures of Association, Contiguous Alternatives, Asymptotic Power, V-statistic

1. Introduction

Generalized partially linear regression models have important applications in *Economics*, *Biological Sciences*, *Environmental science*, *Business administration* etc. A generalized partially linear regression has several utilities in the literature of Statistics. When some of the concomitants jointly display some approximate linear dependence on the study variable of interest or response, it is generally recommended to carry on with such regression model. Here, the study variable is generally explained through the linear function $(\beta_1 X_1 + \ldots + \beta_p X_p)$ of the parametric regressors $X_1, \ldots, X_p$ with unknown constants $\beta_1, \ldots, \beta_p$ and an intractable nonparametric function $m(W_1, \ldots, W_q)$ of the q nonparametric regressors $W_1, \ldots, W_q$; $p, q \geq 2$. In such regression setup, an important objective is determination of estimated response $\hat{Y}$ of response Y on the basis of available information on the regressors $X_1, \ldots, X_p$ as well as $W_1, \ldots, W_q$, for further prediction on Y. Many authors in recent past extensively investigated on the utilities of various semiparametric regression models including the partially linear model, the one with remarkable significance. Using suitable parametric and nonparametric estimation methods, one can further obtain reasonable estimators of the response, parameters and the nonparametric regression function. This method was followed by several authors including *Robinson* (1988) [16], *Cuzick* (1992) [5], *Andrews* (1995) [1], *Hamilton* (1997) [10], *Liu et al.* (1997) [12], *Qi Li* (2000) [14], *Wang et al.* (2011) [19] etc. to derive efficient estimators of the model components as well as their asymptotic properties. *Das et al.* (2022) [6] studied the dependence between sole nonparametric covariate and the error in a simple partially linear model, with the aid of nonparametric regression technique for estimation of kernel density due to Nadaraya-Watson to estimate the unknown nonparametric regression function. In addition, *Robinson* (1988) [16]'s method was helpful to find out consistent estimator of order $n^{-1/2}$ of the sole parameter in the model. On the other hand, *Zhou et al.* (2021) [20] studied the asymptotic properties of V-statistic which are quite relevant in this article.

The article is organized as follows. In Section 2, the unknown parameters $\beta_1, \ldots, \beta_p$ are estimated by Robinson's technique and the nonparametric function $m(W_1, \ldots, W_q)$ is estimated by Nadaraya-Watson estimation of kernel density method. Combining these two estimation procedures, the estimation of response Y is furnished completely. Next, the relevant hypotheses of interest to test whether the joint covariates $(X_1, \ldots, X_p, W_1, \ldots, W_q)$ is independent to ϵ are formulated in Section 3, by applying the theory of contiguity established by *Le Cam* (1960) [4]. Section 4 presents the statistics to be used for the purpose of testing independence in this context, which is indeed a sample analogue of *Kendall* (1938) [11]'s τ. The test statistic is developed by the concept of V-statistic based on some paired samples on $(\hat{Y}^*(r), Y^*(r))$ where $\hat{Y}^*(r)$ denotes the r-th order difference of estimated Y and $\hat{Y}^*(r)$ is the r-th order difference of Y itself. Also, the asymptotic distributions of the test statistic under both the null hypothesis and the contiguous alternatives are derived in this section under various relevant assumptions. Following the limiting distributions of the test statistics, Section 5 delineates its power analysis. For further improvement of its asymptotic power, the order of difference r has been increased gradually. The power curves of the test statistics under the setup of contiguous alternatives are generated against increasing order of difference r for depiction of improved power performance of the test statistic by defining some conditional errors with specified distributions.

2. Model Estimation

The p parameters β_1, …, β_p are estimated by the technique provided by *Robinson* (1988) [16]. To estimate the nonparametric regression function $m(\cdot, \ldots, \cdot)$ we consider the usual kernel density estimation method developed by Nadaraya-Watson. Let us re-express the model as follows.

$$Y_i = \underset{\sim}{X}_i^T \underset{\sim}{\beta} + m(\underset{\sim}{W}_i) + \epsilon_i \tag{1}$$

where $\underset{\sim}{X}_i = (X_{i1}, \ldots, X_{ip})^T$, $\underset{\sim}{W}_i = (W_{i1}, \ldots, W_{iq})^T$, $i = 1, \ldots, n$. Taking conditional expectation to both sides of (1) with respect to $\underset{\sim}{W}_i$, the model gets transformed to the following setup.

$$\begin{aligned} E(Y_i|\underset{\sim}{W}_i) &= E(\underset{\sim}{X}_i|\underset{\sim}{W}_i)^T \underset{\sim}{\beta} + m(\underset{\sim}{W}_i) + E(\epsilon_i|\underset{\sim}{W}_i) \\ i.e.\ \ m_Y(\underset{\sim}{w}) &= m_{\underset{\sim}{X}}(\underset{\sim}{w})^T \underset{\sim}{\beta} + m(\underset{\sim}{w}),\ \text{say.} \end{aligned} \tag{2}$$

Subtracting (2) from (1), we get

$$\begin{aligned} Y_i - m_Y(\underset{\sim}{w}) &= (\underset{\sim}{X}_i - m_{\underset{\sim}{X}}(\underset{\sim}{w}))^T \underset{\sim}{\beta} + \epsilon_i \\ \implies \epsilon_{Yi} &= \epsilon_{\underset{\sim}{X}i}^T \underset{\sim}{\beta} + \epsilon_i,\ \text{where} \\ Y_i &= m_Y(\underset{\sim}{W}_i) + \epsilon_{Yi} \qquad (3) \\ \underset{\sim}{X}_i &= m_{\underset{\sim}{X}}(\underset{\sim}{W}_i) + \epsilon_{\underset{\sim}{X}i},\ i = 1, \ldots, n \qquad (4) \\ \implies \underset{\sim}{\beta}^* &= \left(\sum_{i=1}^{n} \epsilon_{\underset{\sim}{X}i} \epsilon_{\underset{\sim}{X}i}^T\right)^{-1} \left(\sum_{i=1}^{n} \epsilon_{\underset{\sim}{X}i} \epsilon_{Yi}\right). \qquad (5) \end{aligned}$$

Due to the presence of $\epsilon_{\underset{\sim}{X}i}$ and ϵ_{Yi}, the estimator $\underset{\sim}{\beta}^*$ becomes an infeasible quantity. So, it has less feasibility indeed and its improvement can be made by estimating the two error quantities as follows.

From models (3) and (4), the estimators of $m_Y(\cdot)$ and $m_{\underset{\sim}{X}}(\cdot)$ are yielded by applying Nadaraya-Watson estimation of kernel density method, as provided next. Note that,

$$\begin{aligned} m_Y(\underset{\sim}{w}) &= E(Y|\underset{\sim}{W} = \underset{\sim}{w}) \qquad (6) \\ &= \int_{-\infty}^{\infty} y \cdot f_{Y|\underset{\sim}{W}}(y|\underset{\sim}{w})\ dy \qquad (7) \\ &= \int_{-\infty}^{\infty} y \cdot \frac{P_{Y,\underset{\sim}{W}}(y, \underset{\sim}{w})}{g_{\underset{\sim}{W}}(\underset{\sim}{w})}\ dy \qquad (8) \end{aligned}$$

where $\underset{\sim}{w} = (w_1, \ldots, w_q)'$, $P_{Y,\underset{\sim}{W}}(\cdot, \cdot)$ is the joint $p.d.f.$ of $(Y, \underset{\sim}{W})$ and $g_{\underset{\sim}{W}}(\cdot)$ is the $p.d.f.$ of $\underset{\sim}{W}$. The kernel density of $\underset{\sim}{W}$ is estimated as

$$\hat{g}_{\underset{\sim}{W}}(\underset{\sim}{w}) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_1 \ldots h_q}\ k\left(\frac{w_1 - W_{i1}}{h_1}, \ldots, \frac{w_q - W_{iq}}{h_q}\right) \tag{9}$$

at $\underset{\sim}{W} = \underset{\sim}{w}$. Here, $k(\cdot, \ldots, \cdot)$ is the q-dimensional kernel density function of $\underset{\sim}{W}$. We can further simplify the expression of $\hat{g}_{\underset{\sim}{W}}(\cdot)$ as

$$\hat{g}_{\underset{\sim}{W}}(\underset{\sim}{w}) = \frac{1}{n} \sum_{i=1}^{n} \left\{ \prod_{j=1}^{q} \frac{1}{h_j}\ k_j\left(\frac{w_j - W_{ij}}{h_j}\right) \right\} \tag{10}$$

where $k_j(\cdot)$'s are the kernel density functions of W_j's, $j = 1, \ldots, q$; $h_1, \ldots, h_q$ are the bandwidths (> 0) for estimation of kernel density functions of $W_1, \ldots, W_q$. In similar manner, the joint $p.d.f.$ of $(Y, \underset{\sim}{W})$ is estimated as

$$\hat{P}_{Y,\underset{\sim}{W}}(y,\underset{\sim}{w}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_y}k_y\left(\frac{y-Y_i}{h_y}\right)\Big\{\prod_{j=1}^{q}\frac{1}{h_j}k_j\left(\frac{w_j-W_{ij}}{h_j}\right)\Big\}$$

where h_y is the bandwidth for estimating the $p.d.f.$ of Y. Then, at $\underset{\sim}{W}=\underset{\sim}{w}$, $\hat{m}_Y(\underset{\sim}{w})$ is estimated as

$$\hat{m}_Y(\underset{\sim}{w}) = \int_{-\infty}^{\infty} y\cdot\frac{\hat{h}_{Y,\underset{\sim}{W}}(y,\underset{\sim}{w})}{\hat{g}_{\underset{\sim}{W}}(\underset{\sim}{w})}\,dy = \frac{\frac{1}{n}\sum_{i=1}^{n}\Big\{\prod_{j=1}^{q}\frac{1}{h_j}k_j\left(\frac{w_j-W_{ij}}{h_j}\right)\Big\}}{\frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{q}\frac{1}{h_j}k_j\left(\frac{w_j-W_{ij}}{h_j}\right)}.$$

Next, observe that

$$m_{\underset{\sim}{X}}(\underset{\sim}{w}) = E(\underset{\sim}{X}|\underset{\sim}{W}=\underset{\sim}{w}) = \int_{\mathbb{R}^p}\underset{\sim}{x}\cdot g_{\underset{\sim}{X}|\underset{\sim}{W}}(\underset{\sim}{x}|\underset{\sim}{w})\,d\underset{\sim}{x} = \int_{\mathbb{R}^p}\underset{\sim}{x}\cdot\frac{H_{\underset{\sim}{X},\underset{\sim}{W}}(\underset{\sim}{x},\underset{\sim}{w})}{g_{\underset{\sim}{W}}(\underset{\sim}{w})}\,d\underset{\sim}{x}$$

where $H_{\underset{\sim}{X},\underset{\sim}{W}}(\cdot,\cdot)$ is the joint *pdf* of $(\underset{\sim}{X},\underset{\sim}{W})$ estimated as

$$\hat{H}_{\underset{\sim}{X},\underset{\sim}{w}}(\underset{\sim}{x},\underset{\sim}{w}) = \frac{1}{n}\sum_{i=1}^{n}\prod_{m=1}^{p}\frac{1}{a_m}k_{m;X_i}\left(\frac{x_m-X_{im}}{a_m}\right)\Big\{\prod_{z=1}^{q}\frac{1}{b_z}k_{z;W_i}\left(\frac{w_z-W_{iz}}{b_z}\right)\Big\}$$

where $a_1,\dots,a_p$ are the bandwidths for estimating the kernel densities of $X_1,\dots,X_p$ and $b_1,\dots,b_q$ are the bandwidths for estimating the kernel densities of $W_1,\dots,W_q$. Next, $m_{\underset{\sim}{X}}(\underset{\sim}{w})$ is estimated as

$$\hat{m}_{\underset{\sim}{X}}(\underset{\sim}{w}) = \frac{\frac{1}{n}\sum_{i=1}^{n}\Big\{\prod_{m=1}^{p}X_{im}\Big\}\Big\{\prod_{z=1}^{q}\frac{1}{b_z}k_{z;W_i}\left(\frac{w_z-W_{iz}}{b_z}\right)\Big\}}{\frac{1}{n}\sum_{i=1}^{n}\Big\{\prod_{z=1}^{q}\frac{1}{b_z}k_{z;W_i}\left(\frac{w_z-W_{iz}}{b_z}\right)\Big\}}.$$

After estimating $\hat{m}_Y(\cdot)$ and $\hat{m}_{\underset{\sim}{X}}(\cdot)$, the errors are estimated further as $\hat{\epsilon}_{Yi}=Y_i-\hat{m}_Y(\underset{\sim}{W}_i)$ and $\hat{\epsilon}_{\underset{\sim}{X}i}=\underset{\sim}{X}_i-\hat{m}_{\underset{\sim}{X}}(\underset{\sim}{W}_i)$, followed by feasible estimation of $\underset{\sim}{\beta}$ as $\hat{\underset{\sim}{\beta}}=\left(\sum_{i=1}^{n}\hat{\epsilon}_{\underset{\sim}{X}i}\hat{\epsilon}_{\underset{\sim}{X}i}^T\right)^{-1}\left(\sum_{i=1}^{n}\hat{\epsilon}_{\underset{\sim}{X}i}\hat{\epsilon}_{Yi}\right)$. Next, our objective is to estimate the nonparametric regression function $m(\cdot,\dots,\cdot)$. Note that, the partially linear model can be further transformed to a nonparametric regression model as $Y_i-\underset{\sim}{X}_i^T\underset{\sim}{\beta}=m(\underset{\sim}{W}_i)+\epsilon_i \implies Y_i^{'}=m(\underset{\sim}{W}_i)+\epsilon_i,\ i=1,\dots,n$ where $Y_i^{'}=Y_i-\underset{\sim}{X}_i^T\underset{\sim}{\beta}$ is the *transformed response* for all $i=1,\dots,n$. Based on $i.i.d.$ observations $(Y_i^{'},\underset{\sim}{W}_i)$' s, $i=1,\dots,n$, $m(\cdot)$ is finally estimated as

$$\hat{m}(\underset{\sim}{w}) = \frac{\frac{1}{n}\sum_{i=1}^{n}\Big\{\prod_{j=1}^{q}k_{j;W_i}\left(\frac{w_j-W_{ij}}{h_j}\right)\Big\}Y_i^{'}}{\frac{1}{n}\sum_{i=1}^{n}\Big\{\prod_{j=1}^{q}k_{j;W_i}\left(\frac{w_j-W_{ij}}{h_j}\right)\Big\}},$$

where $\hat{Y}_i^{'}=(Y_i-\underset{\sim}{X}_i^T\hat{\underset{\sim}{\beta}})$.

Here, the Nadaraya-Watson kernel density estimation method has been preferred over any other traditional nonparametric smoothing method like spline smoothing or polynomial smoothing etc., solely because a semiparametric regression model is expected to have lower dimensionality compared to a traditional multiple nonparametric regression model. The Nadaraya-Watson kernel density estimation technique provides more accurate estimated value of the nonparametric regression function in a semiparametric model whereas spline smoothing, polynomial smoothing are optimum choices in the context of a nonparametric regression model. For the choices of bandwidths, the idea of *Silverman* (2018) [17] is needed in this discourse.

3. Hypotheses

The independence between jointly distributed regressors $X_1,\dots,X_p,W_1,\dots,W_q$ and error ϵ could be a matter of investigation in our proposed partially linear regression setup. The natural hypotheses of interest in this regard are $H_0:\underset{\sim}{Z}\perp\!\!\!\perp\epsilon$ and $H_1:\underset{\sim}{Z}\not\perp\!\!\!\perp\epsilon$ where $\underset{\sim}{Z}=(\underset{\sim}{X},\underset{\sim}{W})$ is the

$(p+q)$-dimensional vector of regressors. But we have to work on the hypotheses further due to unobservability of ϵ by defining a suitable function of ϵ on $\mathbb{R}$, say $s(\epsilon)$, which is indeed tractable.

Instead of single error ϵ, we would like to consider a function of several $i.i.d.$ errors, say $s(\epsilon_1,\ldots,\epsilon_t)$, $t\in\mathbb{N}$. The proposition of $s(\epsilon_1,\ldots,\epsilon_t)$ was made by various authors in recent past. Einmahl et al. (2008) [9] defined a second order difference of three $i.i.d.$ errors as $(\epsilon_1-2\epsilon_2+\epsilon_3)$ to conduct a test of hypothesis of independence between X and ϵ under a simple nonparametric regression model $Y=m(X)+\epsilon$. The function $(\epsilon_1-2\epsilon_2+\epsilon_3)$ is well approximated by $(Y_1-2Y_2+Y_3)$ due to smoothness assumption of the nonparametric regression function of $m(\cdot)$ where Y_1,Y_2,Y_3 are three observations on response Y corresponding to $\epsilon_1,\epsilon_2,\epsilon_3$. Such assumption is required solely because of the unobservability of $\epsilon_1,\epsilon_2,\epsilon_3$. Later, an extended idea was executed by Das et al. (2023) [7], by taking third order of difference of three $i.i.d.$ errors $(\epsilon_1-3\epsilon_2+3\epsilon_3-\epsilon_4)$ under the same nonparametric regression context based on some nonparametric measures of association based on Kendall's τ, Bergsma et al. (2014)'s τ^* and Szekely et al. (2007) [18]'s $dCov$, which indeed contributes more than the corresponding second order difference of ϵ in achieving more powerful test of independence between X and ϵ.

In addition, taking a simple partially linear regression model $Y=Z\beta+m(X)+\epsilon$, Das et al. (2022) [6] tested the independence between only the parametric regression coefficient X and the error ϵ and established the power of the tests based on τ, τ^* and $dCov$ where they defined the second order difference of ϵ, later approximated by the corresponding second order difference of response values. Hence, there might be an additional question regarding the performance of the test statistics if one proceeds to check the independence between (Z,X) and ϵ under the same model. In this work, a more generalized version of a partially linear regression model has been taken, and the objective lies in testing independence between all the regressors and error. Moreover, the testing scheme would involve a more general order difference of ϵ (or, equivalently, Y as $m(\cdot,\ldots,\cdot)$ is assumed to be Lipschitz continuous here) rather than specified ones like second or third order differences, as mentioned above.

Hence, a real-valued function of the $(p+q)$ regressors is needed to perform the test, and H_0 further implies that the function is independent of $s(\epsilon_1,\ldots,\epsilon_t)$ where $\epsilon_1,\ldots,\epsilon_t$ are t $i.i.d.$ errors. To obtain a well-specified null hypothesis, we first define a general r-th order differences of ϵ and Y as

$$\epsilon^*(r) = \sum_{j=1}^{r+1}(-1)^{j-1}\binom{r}{j-1}\epsilon_j$$

and

$$Y^*(r) = \sum_{j=1}^{r+1}(-1)^{j-1}\binom{r}{j-1}Y_j$$

where $\epsilon_1,\ldots,\epsilon_{r+1}$ and $Y_1,\ldots,Y_{r+1}$ are $(r+1)$ $i.i.d.$ errors and responses, respectively. First, we would like to discuss why considering the r-th order difference of error is better than any general linear function of the errors.

Proposition 3.1. $\epsilon^*(r)$ has maximal k-th order absolute moment among all possible linear functions $\sum_{j=1}^{r+1}u_j\epsilon_j$ with real coefficients u_j's

The above proposition holds for $r=2$ as delineated by Dhar et al. (2018) [8] $r=3$ as deduced by Das et al. (2023) [7]. The general proof is available in Appendix-I.

Theorem 3.1. $Y^*(r)\approx\epsilon^*(r)$. Moreover, among all linear functions, $\epsilon^*(r)$ has the maximum k-th order absolute moment.

The above theorem is quite meaningful in this context. It is evident that $\epsilon^*(1)$ can be well approximated by $Y^*(1)$, but one can obtain a test of homoscedasticity of ϵ based on $\epsilon^*(1)$. Hence, *Das et al.*,(2023) [7] considered $\epsilon^*(2)$ and approximated it by $Y^*(2)$ to construct a test scheme for the independence between nonparametric covariate X and error ϵ. Furthermore, in a simple partially linear regression setup $Y=Z\beta+m(X)+\epsilon$, *Das et al.* (2022) [6] considered $\epsilon^*(3)$ to develop a test of independence between the sole nonparametric covariate X and the random error ϵ. The proof of Theorem 3.1 is elaborately discussed in *Appendix-II* to understand the consideration of general order difference of ϵ in this regard.

Finally, we get transformed null hypothesis as $H_0:\ (\underset{\sim}{X},\underset{\sim}{W})\perp\!\!\!\perp\epsilon^*(r)$ which further implies approximately that $(\underset{\sim}{X},\underset{\sim}{W})\perp\!\!\!\perp Y^*(r)$. Again, as another implication of H_0, any continuous function of $(\underset{\sim}{X},\underset{\sim}{W})$ is independent to $Y^*(r)$. The following proposition is helpful to proceed with $\hat{Y}$ in this testing scheme. Proof is available in *Appendix-II*.

Proposition 3.2. $\hat{Y}^*(r)$ can be approximated as a function of $(\underset{\sim}{X},\underset{\sim}{W})$ where $\hat{Y}^*(r)$ is the r-th order difference of $\hat{Y}$.

Therefore, the null hypothesis further implies that $H_0:\ \hat{Y}^*(r)\perp\!\!\!\perp Y^*(r)$ and subsequently the hypotheses of interest are derived as

$$H_0:\ \hat{Y}^*(r)\perp\!\!\!\perp Y^*(r)\ \text{ against }\ H_1:\ \hat{Y}^*(r)\not\perp\!\!\!\perp Y^*(r). \tag{11}$$

(11) further implies that $H_0:\mathcal{M}(\hat{Y}^*(r),Y^*(r))=0$ against $H_1:\mathcal{M}(\hat{Y}^*(r),Y^*(r))\neq 0$, where $\mathcal{M}$ is the concerning nonparametric measure of association between $\hat{Y}^*(r)$ and $Y^*(r)$. A more formal approach for testing the hypotheses is to develop a consistent test procedure. An useful wayout to perform the test consistently is forming a sequence of contiguous alternative hypotheses due to *Le Cam* (1960) [4]. Such a sequence converges to H_0 with increasing sample size. The following sequence of alternative hypotheses (see *Das et al.* (2023) [7]), a contiguous one, is considered as

$$H_n : F_{n;\hat{Y}^*(r),Y^*(r)}(\hat{y}^*(r), y^*(r)) = \left(1 - \frac{\mu}{\sqrt{n}}\right) F_{0;\hat{Y}^*(r),Y^*(r)}(\hat{y}^*(r), y^*(r)) + \frac{\mu}{\sqrt{n}} F_{\hat{Y}^*(r),Y^*(r)}(\hat{y}^*(r), y^*(r)) \quad (12)$$

where $F_{n;\hat{Y}^*(r),Y^*(r)}(\cdot,\cdot)$ is the joint CDF of $(\hat{Y}^*(r), Y^*(r))$ under H_n, $F_{0;\hat{Y}^*(r),Y^*(r)}(\cdot,\cdot)$ and $F_{\hat{Y}^*(r),Y^*(r)}(\cdot,\cdot)$ are the joint CDFs of $(\hat{Y}^*(r), Y^*(r))$ under H_0 and H_1 respectively, $\mu(> 0)$ is the tuning parameter.

Theorem 3.2. Under three following assumptions:

1. $\frac{\partial^2}{\partial \hat{y}^* \partial y^*} F_0(\hat{y}^*(r), y^*(r)) < \infty$, say $f_0(\hat{y}^*(r), y^*(r)) \;\forall\; (\hat{y}^*(r), y^*(r))$,
2. $\frac{\partial^2}{\partial \hat{y}^* \partial y^*} F(\hat{y}^*(r), y^*(r)) < \infty$, say $f(\hat{y}^*(r), y^*(r)) \;\forall\; (\hat{y}^*(r), y^*(r))$,
3. $E_{F_0}\left[\frac{f(\hat{y}^*(r), y^*(r))}{f_0(\hat{y}^*(r), y^*(r))} - 1\right]^2 < \infty$,

H_n is a contiguous sequence of alternative hypotheses.

The proof is furnished by *Das et al.* (2022) [6, p. 559].

Here $f_0(\cdot,\cdot)$ and $f(\cdot,\cdot)$ are the joint probability density function of $(\hat{Y}^*(r), Y^*(r))$ under H_0 and H_1 respectively. Next, we proceed to test H_0 against H_n, by constructing suitable test statistics based on the theory of non-degenerate V-statistic.

4. Test Statistics

To test the null hypothesis that $\hat{Y}^*(r)$ is independent to $Y^*(r)$, a convenient approach is to develop robust test statistics. With the changing values of r, the asymptotic properties of the test statistics are expected to vary. For instance, one may study on the asymptotic power analysis of the test statistics for different choices of r. A frequently utilised measure of association, namely Kendall's τ, is considered in this backdrop to evaluate power performance of the test statistics made upon τ. A s discussed earlier, we proceed to test

$$H_0 : \tau(\hat{Y}^*(r), Y^*(r)) = 0 \;\text{ vs. }\; H_1 : \tau(\hat{Y}^*(r), Y^*(r)) \neq 0.$$

To test the above hypotheses, we construct appropriate V-statistic based on paired observations from the bivariate CDF of $(\hat{Y}^*(r), Y^*(r))$. Let $(\hat{y}_1^*, y_1^*), \ldots, (\hat{y}_n^*, y_n^*)$ be the n i.i.d. samples from the joint CDF. Then, the kernel of Kendall's τ is defined as

$$\begin{aligned} & h((\hat{y}^*_{\alpha_1}(r), y^*_{\alpha_1}(r)), (\hat{y}^*_{\alpha_2}(r), y^*_{\alpha_2}(r))) \quad (13)\\ = \;& sign\{(\hat{y}^*_{\alpha_1}(r) - \hat{y}^*_{\alpha_2}(r))(y^*_{\alpha_1}(r) - y^*_{\alpha_2}(r))\}. \quad (14)\end{aligned}$$

for $1 \leq \alpha_1 \neq \alpha_2 \leq n$. The corresponding V-statistic to test H_0 against H_n is regarded as the test statistic of interest in our discussion. The form of the test statistic is provided as

$$T_n^{(r)} = \frac{1}{n^2} \sum_{\alpha_1,\alpha_2=1}^{n} sign\{(\hat{y}^*_{\alpha_1}(r) - \hat{y}^*_{\alpha_2}(r))(y^*_{\alpha_1}(r) - y^*_{\alpha_2}(r))\}.$$

Since $h((\hat{y}^*_{\alpha_1}(r), y^*_{\alpha_1}(r)), (\hat{y}^*_{\alpha_2}(r), y^*_{\alpha_2}(r)))$ has order of degeneracy 0 (*Das et al.* (2022) [6]), $T_n^{(r)}$ is a nondegenerate V-statistic indeed. Then, the asymptotic distribution of $\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)}))$ is a Gaussian one with zero expectation and a constant variance. The limiting distributions of $T_n^{(r)}$, both under H_0 and H_n, are required to determine its asymptotic power for different values of r. *Zhou et al.* (2021) [20] discussed on the asymptotic normality of a nondegenerate V-statistic in detail.

Theorem 4.1. Under H_0,
$\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) \xrightarrow{L} N(0, 4\eta_{1,2}(r))$, provided that $E[h^2((\hat{Y}_1^*(r), Y_1^*(r)), (\hat{Y}_2^*(r), Y_2^*(r)))] < \infty$ with

$$\eta_{1,2}(r) = Var[E(h((\hat{Y}_1^*(r), Y_1^*(r)), (\hat{Y}_2^*(r), Y_2^*(r)))\Big|(\hat{Y}_1^*(r), Y_1^*(r)))] = \frac{1}{9}. \quad (15)$$

Theorem 4.2. Under H_n,

$$\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) \xrightarrow{L} N(\nu^{(r)}, 4\eta_{1,2}(r)),$$

where

$$\nu^{(r)} = \lim_{n\to\infty} Cov_{H_0}\left(\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)})), \log \frac{dF_n}{dF_0}\right).$$

The theorem is similar to Theorem 9 (iv) of *Das et al. (2022)*[6]. Le Cam's third lemma is useful to complete the proof. Also, $\nu^{(r)} = \mu(2P_{c1} - 1)$ where P_{c1} is the probability of concordance of n *i.i.d.* samples from $(\hat{Y}^*(r), Y^*(r))$. Furthermore, $\nu^{(r)}$

can also be deduced as

$$\mu\, E_{H_0}\left[h((\hat{Y}_1^*(r), Y_1^*(r)), (\hat{Y}_2^*(r), Y_2^*(r))) \times \frac{f(\hat{Y}_1^*(r), Y_1^*(r))}{f_0(\hat{Y}_1^*(r), Y_1^*(r))}\right].$$

Next, we compute asymptotic power of $\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)}))$ in the proposed test setup.

5. Asymptotic Powers

The statistical power of a test statistic determines how it behaves when the null of hypothesis lacks its rationality. A reasonably powerful test statistic is always desired to perform a given test of hypothesis. However, consistency plays a key role to delineate the utility of the test statistic. In this circumstance, the setup of contiguous alternatives is considered only to develop a consistent test statistic $T_n^{(r)}$.

The asymptotic power of $T_n^{(r)}$ is determined as $P_{H_n}(\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) > t_\kappa) = 1 - \Phi\left(\frac{t_\kappa - \nu^{(r)}}{\sqrt{4\eta_{1,2}(r)}}\right)$, where t_κ satisfies $P_{H_0}(\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) > t_\kappa) = \kappa$ with $0 < \kappa < 1$ and $\Phi(\cdot)$ is the CDF of a standard normal distribution. For $\mu = 0$, $\nu^{(r)} = 0$ for which asymptotic power and size of $T_n^{(r)}$ are equal. Otherwise, as $\mu \uparrow$, the limiting power of $T_n^{(r)}$ increases, provided the probability of concordance exceeds 0.5.

Next, the consistency of $\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)}))$ as well as its bigger power for increasing sample size need to be checked to perform the relevant power study eventually.

Proposition 5.1. For $n^* > n$, $P_{H_n}(\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) > t_\kappa) < P_{H_n}(\sqrt{n^*}(T_{n^*}^{(r)} - E(T_{n^*}^{(r)})) > t_\kappa)$ and $P_{H_n}(\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) > t_\kappa) \uparrow 1$ as $\mu \uparrow$ and $n \to \infty$.

Now we shall consider various examples to study the power performance of $T_n^{(r)}$ against the tuning parameter μ by taking the sample size $n = 1000$ and the orders of difference $r = 2, 3, 4, 5, 10$.

5.1. Examples

1. We consider a generalized partially linear model $Y = \beta_1 X_1 + \beta_2 X_2 + m(W_1, W_2) + \epsilon$ with usual assumptions on error ϵ, $viz.$, (i) $E(\epsilon | X_1 = x_1, X_2 = x_2, W_1 = w_1, W_2 = w_2) = 0$ and (ii) $E(\epsilon^2 | X_1 = x_1, X_2 = x_2, W_1 = w_1, W_2 = w_2) = \sigma^2(x_1, x_2, w_1, w_2)$ for all (x_1, x_2, w_1, w_2). The joint distribution of $\underset{\sim}{Z} = (X_1, X_2, W_1, W_2)^T$ is $\mathbb{N}_4\left(\begin{pmatrix}0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0.18 & -0.06 & 0.22 & -0.13\\ -0.06 & 0.14 & -0.28 & 0.19\\ 0.22 & -0.28 & 0.2 & 0.17\\ -0.13 & 0.19 & 0.17 & 0.25\end{pmatrix}\right)$, which is independent of $\epsilon \sim N(0, 0.015)$ under H_0. In addition, Y has a t-distribution with 2 degrees of freedom. The nonparametric regression function is considered as $m(W_1, W_2) = 0.45 W_1 W_2 - 0.25 W_1^2 W_2 + W_2^3$. The conditional error distributions for $\underset{\sim 1}{c} = (-1.5, -1.7, 1.2, 1.3)' \in \mathbb{R}^4$ under H_1 are

Example 5.1. $\epsilon \,\big|\, \underset{\sim}{Z} \sim N(0,\, 0.015\, |1 + \underset{\sim 1}{c}^T \underset{\sim}{Z}|)$.

Example 5.2. $\epsilon \,\big|\, \underset{\sim}{Z} \overset{\mathcal{D}}{=} \frac{\chi^2_{\underset{\sim}{Z}} - d_{\underset{\sim}{Z}}}{\sqrt{2\, d_{\underset{\sim}{Z}}}}$ where $d_{\underset{\sim}{Z}} = |1 + \underset{\sim 1}{c}^T \underset{\sim}{Z}|^{-1}$ and $\chi^2_{\underset{\sim}{Z}} \sim \chi^2_{\lceil d_{\underset{\sim}{Z}} \rceil}$.

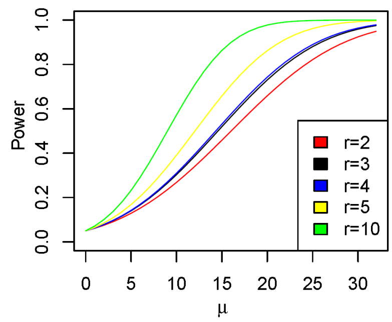

Figure 1. *Asymptotic powers of $T_n^{(r)}$ in Example 5.1 where $p = q = 2$.*

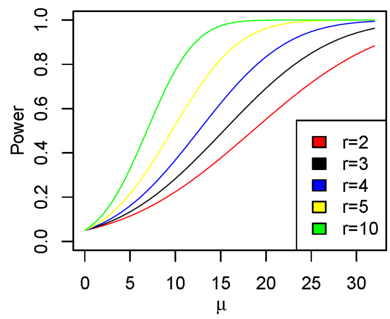

Figure 2. *Asymptotic powers of $T_n^{(r)}$ in Example 5.2 where $p = q = 2$.*

2. A generalized partially linear model $Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + m(W_1, W_2) + \epsilon$ is considered along with assumptions on ϵ such as (i) $E(\epsilon | X_1 = x_1, X_2 = x_2, X_3 = x_3, W_1 = w_1, W_2 = w_2) = 0$ and (ii) $E(\epsilon^2 | X_1 = x_1, X_2 = x_2, X_3 = x_3, W_1 = w_1, W_2 = w_2) = \sigma^2(x_1, x_2, x_3, w_1, w_2)$ for all $(x_1, x_2, x_3, w_1, w_2)$. The joint distribution of $\underset{\sim}{Z} = (X_1, X_2, X_3, W_1, W_2)^T$ is $\mathbb{N}_5\left(\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0.18 & -0.06 & 0.09 & 0.32 & -0.18\\ -0.06 & 0.14 & -0.17 & 0.25 & 0.15\\ 0.09 & -0.17 & 0.25 & -0.3 & -0.23\\ 0.32 & 0.25 & -0.3 & 0.2 & 0.17\\ -0.18 & 0.15 & -0.23 & 0.17 & 0.25\end{pmatrix}\right)$,

which is independently distributed to $\epsilon \sim N(0, 0.015)$ under H_0. In addition, $Y \sim t_2$. The nonparametric regression function is considered as $m(W_1, W_2) = 0.45W_1W_2 - 0.25W_1^2W_2 + W_2^3$. Under H_1 for $\underset{\sim 2}{c} = (-1.5, -1.7, 1.2, 1.3, -3.5)' \in \mathbb{R}^5$, the conditional error distributions are considered as

Example 5.3. Under H_1, $\epsilon \Big| \underset{\sim}{Z} \sim N(0,\, 0.015\, |1 + \underset{\sim 2}{c}^T \underset{\sim}{Z}|)$

Example 5.4. Under H_1, $\epsilon \Big| \underset{\sim}{Z} \overset{\mathcal{D}}{=} \dfrac{\chi^2_{\underset{\sim}{Z}} - d'_{\underset{\sim}{Z}}}{\sqrt{2\, d'_{\underset{\sim}{Z}}}}$ where $d'_{\underset{\sim}{Z}} = |1 + \underset{\sim 2}{c}^T \underset{\sim}{Z}|^{-1}$ and $\chi^2_{\underset{\sim}{Z}} \sim \chi^2_{\lceil d'_{\underset{\sim}{Z}} \rceil}$.

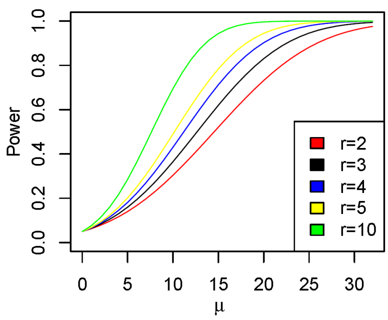

Figure 3. *Asymptotic powers of $T_n^{(r)}$ in Example 5.3 where $p = 3,\ q = 2$.*

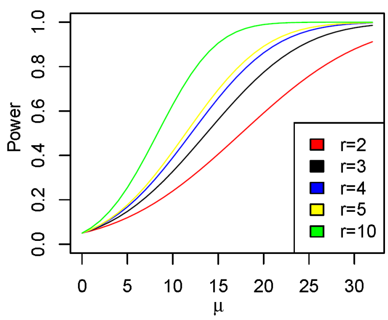

Figure 4. *Asymptotic powers of $T_n^{(r)}$ in Example 5.4 where $p = 3,\ q = 2$.*

3. The generalized partially linear model $Y = \beta_1 X_1 + \beta_2 X_2 + m(W_1, W_2, W_3) + \epsilon$ is considered with assumptions on ϵ as (i) $E(\epsilon | X_1 = x_1, X_2 = x_2, W_1 = w_1, W_2 = w_2, W_3 = w_3) = 0$ and (ii) $E(\epsilon^2 | X_1 = x_1, X_2 = x_2, W_1 = w_1, W_2 = w_2, W_3 = w_3) = \sigma^2(x_1, x_2, w_1, w_2, w_3)$ for all $(x_1, x_2, w_1, w_2, w_3)$. The joint distribution of $\underset{\sim}{Z} = (X_1, X_2, W_1, W_2, W_3)^T$ is
$$\mathbb{N}_5\left(\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix} 0.18 & -0.06 & 0.32 & -0.18 & -0.24 \\ -0.06 & 0.14 & 0.25 & 0.15 & -0.18 \\ 0.32 & 0.25 & 0.2 & 0.17 & -0.22 \\ -0.18 & 0.15 & 0.17 & 0.25 & 0.11 \\ -0.24 & -0.18 & -0.22 & 0.11 & 0.27 \end{pmatrix}\right)$$
which is independently distributed to $e \sim N(0, 0.015)$ under H_0 and $Y \sim t_2$. The nonparametric regression function is considered as $\phi(W_1, W_2, W_3) = 0.45W_1W_2 - 0.25W_1^2W_3 + W_3^3$. Under H_1 for $\underset{\sim 3}{c} = (2.5, 4, -6.2, -5.5, 3.1)' \in \mathbb{R}^5$, the conditional error distributions are considered as

Example 5.5. Under H_1, $\epsilon \Big| \underset{\sim}{Z} \sim N(0,\, 0.015\, |1 + \underset{\sim 3}{c}^T \underset{\sim}{Z}|)$

Example 5.6. Under H_1, $e \Big| \underset{\sim}{Z} \overset{\mathcal{D}}{=} \dfrac{\chi^2_{\underset{\sim}{Z}} - d'_{\underset{\sim}{Z}}}{\sqrt{2\, d'_{\underset{\sim}{Z}}}}$ where $d'_{\underset{\sim}{Z}} = |1 + \underset{\sim 3}{c}^T \underset{\sim}{Z}|^{-1}$ and $\chi^2_{\underset{\sim}{Z}} \sim \chi^2_{\lceil d'_{\underset{\sim}{Z}} \rceil}$.

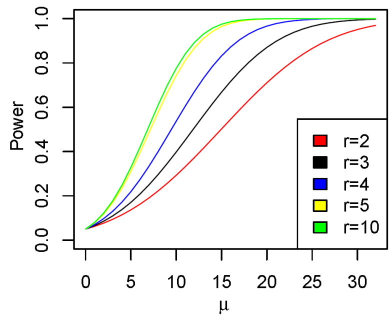

Figure 5. *Asymptotic powers of $T_n^{(r)}$ in Example 5.5 where $p = 2,\ q = 3$.*

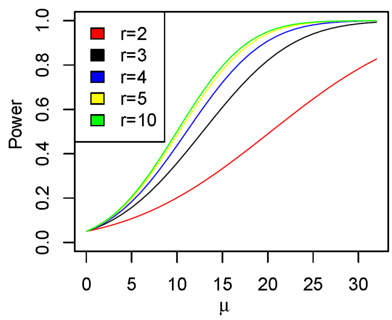

Figure 6. *Asymptotic powers of $T_n^{(r)}$ in Example 5.6 where $p = 2,\ q = 3$.*

4. The generalized partially linear model $Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + m(W_1, W_2, W_3) + \epsilon$ is considered with assumptions on ϵ as (i) $E(\epsilon | X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5, W_1 = w_1, W_2 = w_2, W_3 = w_3) = 0$ and (ii) $E(\epsilon^2 | X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5, W_1 = w_1, W_2 = w_2, W_3 = w_3) = \sigma^2(x_1, x_2, x_3, x_4, x_5, w_1, w_2, w_3)$ for all $(x_1, x_2, x_3, x_4, x_5, w_1, w_2, w_3)$. The joint distribution of $(X_1, X_2, X_3, X_4, X_5, W_1, W_2, W_3)^T$ is $\mathbb{N}_8\left(\begin{pmatrix}\underset{\sim 5}{0}\\ \underset{\sim 3}{0}\end{pmatrix}, \begin{pmatrix}\Sigma_5 & \Sigma_{53} \\ \Sigma_{35} & \Sigma_3\end{pmatrix}\right)$
where

$$\Sigma_5 = \begin{pmatrix} 0.18 & -0.06 & 0.09 & -0.13 & 0.16 \\ -0.06 & 0.14 & -0.17 & 0.26 & -0.14 \\ 0.09 & -0.17 & 0.25 & 0.33 & -0.18 \\ -0.13 & 0.26 & 0.33 & 0.32 & 0.05 \\ 0.16 & -0.14 & -0.18 & 0.05 & 0.24 \end{pmatrix},$$

$$\Sigma_3 = \begin{pmatrix} 0.11 & -0.21 & -0.19 \\ -0.21 & 0.27 & 0.14 \\ -0.19 & 0.14 & 0.4 \end{pmatrix}$$

$$\Sigma_{53} = \begin{pmatrix} -0.17 & 0.31 & -0.22 \\ 0.18 & 0.25 & -0.33 \\ -0.24 & -0.15 & 0.15 \\ -0.14 & -0.07 & 0.12 \\ 0.26 & -0.18 & -0.03 \end{pmatrix},$$

$$\Sigma_{35} = \Sigma_{53}^T = \begin{pmatrix} -0.17 & 0.18 & -0.24 & -0.14 & 0.26 \\ 0.31 & 0.25 & -0.15 & -0.07 & -0.18 \\ -0.22 & -0.33 & 0.15 & 0.12 & -0.03 \end{pmatrix}.$$

Moreover, $\underset{\sim}{Z} = (X_1, X_2, X_3, X_4, X_5, W_1, W_2, W_3)^T$ is distributed independently to $\epsilon \sim N(0, 0.015)$ under H_0 and $Y \sim t_2$. The nonparametric regression function is considered as $m(W_1, W_2, W_3) = 0.36W_1^3 - 0.25W_2^2W_3 - 0.11W_3^2W_1 + 0.08W_1W_2W_3$. The conditional error distributions under H_1 by taking $\underset{\sim 4}{c} = (5.92, -3.78, -10.66, 8.89, -5.45, 9.65, 8.35, -7.89)' \in \mathbb{R}^8$ are considered as

Example 5.7. $\epsilon \,\big|\, \underset{\sim}{Z} \sim N(0,\, 0.015\, |1 + \underset{\sim 4}{c}^T \underset{\sim}{Z}|)$

Example 5.8. $\epsilon \,\big|\, \underset{\sim}{Z} \overset{\mathcal{D}}{=} \dfrac{\chi^2_{\underset{\sim}{Z}} - d'_{\underset{\sim}{Z}}}{\sqrt{2\, d'_{\underset{\sim}{Z}}}}$ where $d'_{\underset{\sim}{Z}} = |1 + \underset{\sim 4}{c}^T \underset{\sim}{Z}|^{-1}$ and $\chi^2_{\underset{\sim}{Z}} \sim \chi^2_{\lceil d'_{\underset{\sim}{Z}} \rceil}$.

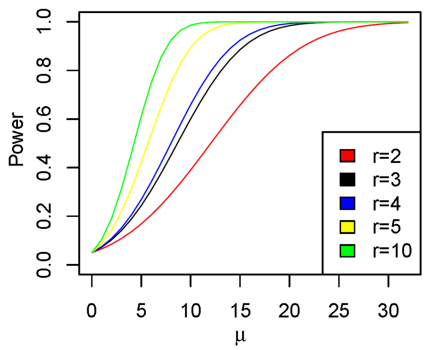

Figure 7. *Asymptotic powers of $T_n^{(r)}$ in Example 5.7 where $p = 5,\ q = 3$.*

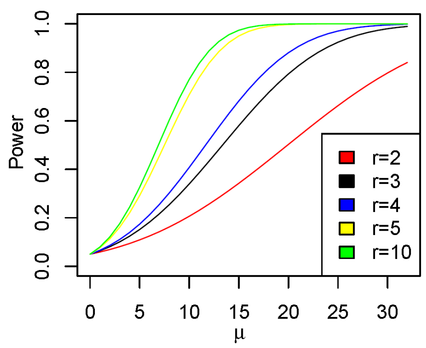

Figure 8. *Asymptotic powers of $T_n^{(r)}$ in Example 5.8 where $p = 5,\ q = 3$.*

One can observe, from the above examples as well as the corresponding power graphs $T_n^{(r)}$'s, that increasing numbers of parametric and nonparametric regressors actually stimulate the asymptotic power performances of the test statistics. It is quite noteworthy that the power curves of $T_n^{(10)}$'s (indicated by green graphs) perform with significantly higher powers in all the examples. Additionally, in case of Gaussian conditional errors involved in the assumed models, the performances of $T_n^{(10)}$'s are better compared to the situations where conditional errors are standardized chi-squared random variables. In few cases, the power curves shown by test statistics $T_n^{(4)}$ and $T_n^{(5)}$ of observed and predicted responses are too close to the power curves of $T_n^{(10)}$. As usual, $T_n^{(2)}$ emerges as lowest performer compared to the other ones, although having power curves being almost same with the power curves of $T_n^{(3)}$ in some cases. Such increment of power performances of $T_n^{(r)}$'s with $r = 2, 3, 4, 5, 10$ further motivates us to consider a test statistic $T_n^{(r)}$ with higher r to conduct the testing of hypotheses with greater consistency.

It is also noteworthy that $T_n^{(r)}$ becomes more powerful if more parametric regressors are incorporated in the model, as shown in Example 5.7 and 5.8 where 5 parametric regressors and 3 nonparametric regressors are present in the model in *Section 5*. It is clearly observed that the power curves of the test statistics are comparatively more rapid than the power curves of them in other instances having two or three parametric regressors considered previously. In general, the increasing number of parametric regressors in a semiparametric regression model gradually transforms the setup to a nonparametric one, as it reduces the model flexibility due to numerous deterministic relationships between response and regressors.

One can express both $\hat{Y}^*(r)$ and $Y^*(r)$ in terms of $(r-1)$ number of first order differences of $\hat{Y}$ and Y-observations respectively. As the order of difference r increases, both of them become linear functions of large number of first order differences of $\hat{Y}$ and Y-observations and those differences are indeed independent to each other. However, there is a genuine dependence between $\hat{Y}$ and Y and the dependence is enhanced for $r \uparrow$. As a result, the power of the concerning test is yielded higher in case of higher order difference situation, which is inevitable for $T_n^{(10)}$ in this context. All the values of asymptotic powers of $T_n^{(r)}$'s are provided in *Appendix I*.

6. Data Analysis

We must study the adequacy of $T_n^{(r)}$'s for real datasets also; hence consider the *Apple quality dataset* (https://www.kaggle.com/datasets/nelgiriyewithana/apple-quality), used to assess the quality of apples. There are 9 variables with 4001 observations each, namely 'A-id' (unique identifier for each apple), 'size' (size of apple), 'weight' (weight of apple), 'sweetness' (degree of sweetness of apple), 'crunchiness' (texture indicating the crunchiness of the apple), ‘juiciness’ (level of juiciness of apple), ‘ripeness’ (stage of ripeness of apple), ‘quality’ (overall quality of apple) and ‘acidity’ (acidity level of apple). All the variables are measured in proper units, except the nominal one ‘A-id’ and ordinal one ‘quality’. The objective is to study the variable ‘acidity’ on the basis of 6 quantitative variables ‘size’, ‘weight’, ‘sweetness’, ‘crunchiness’, ‘juiciness’ and ‘ripeness’.

Generally, any semiparametric regression model (including a partially linear one) involves parametric and nonparametric components where the parametric part captures key structural relationships and the nonparametric part provides flexibility as required. Since we assume some mathematical structure(s) on the underlying data, semiparametric model needs fewer observations compared to fully nonparametric regression models based on fewer assumptions about the functional form of the relationship between variables to achieve the same level of accuracy. In addition, non-parametric regression models are estimated using flexible techniques such as kernel smoothing, spline, or local regression, which can be adapted to complex structures in the data. This flexibility is achieved when we consider a large sample size to obtain reliable estimates and avoid overfitting. Hence, we prefer to reduce the full dataset from 4001 to 150.

Scatterplots of ‘acidity’ versus each of 6 variables are generated for the reduced dataset to observe which variables are explaining ‘acidity’ parametrically and nonparametrically. It is worth observing that both the variables ‘juiciness’ and ‘ripeness’ form deterministic relationships with ‘acidity’; the first one exhibits nearly positive association via approximately linear dependence with positive slope whereas the second one shows downward relationship with soaring level of ‘acidity’ by displaying more or less linear association with negative slope. The association of ‘acidity’ with any of other four variables are not explicit at all, hence there exist nonparametric relationships between the four regressors and the response variable ‘acidity’. Here, the regression model is a partially linear one, expressed as $Y = \beta_1 X_1 + \beta_2 X_2 + m(W_1, W_2, W_3, W_4) + \epsilon$ where X_1 and X_2 denote ‘ripeness’ and ‘juiciness’ respectively, the 4 nonparametric variables W_1, W_2, W_3, W_4 represent ‘size’, ‘crunchiness’, ‘weight’ and ‘sweetness’ respectively and Y denotes ‘acidity’. β_1 and β_2 are estimated as $\hat{\beta}_1 \approx 0.0377$ and $\hat{\beta}_2 \approx -0.0842$. The r-th order difference of observed Y are calculated, followed by evaluations of test statistics for $r = 2, 3, 4, 5, 10$. Furthermore, bootstrap samples are generated on the r-th order differences of $(Y, \hat{Y})$ and the test statistics are computed at all stages of resampling. The p-values of $T_n^{(r)}$’s are estimated that provide the extent of rejection of H_0 at chosen level of significance α $(0 < \alpha < 1)$. Since the asymptotic power of $T_n^{(r)}$ increases as $r \uparrow$, it is expected that the p-values would decrease henceforth.

Next, we present the scatterplots of acidity versus all the covariates under consideration, as well as the p-values of $T_n^{(r)}$’s for different resampling sizes (B) for $r = 2, 3, 4, 5, 10$.

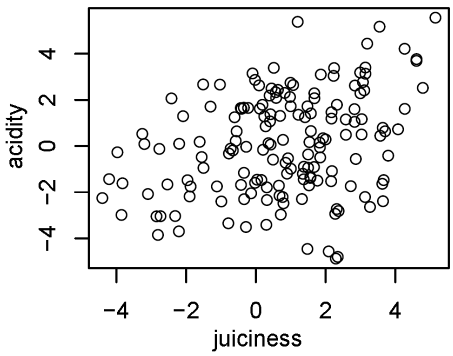

Figure 9. *Scatterplot of acidity vs juiciness.*

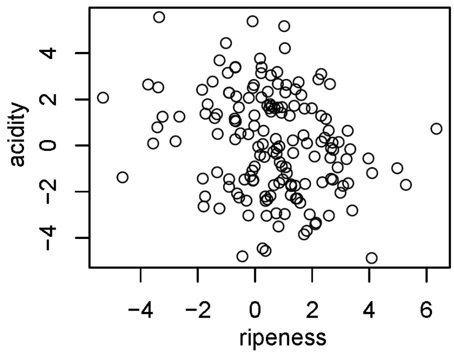

Figure 10. *Scatterplot of acidity vs juiciness as well as ripeness.*

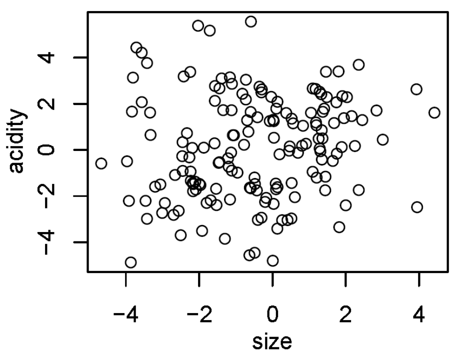

Figure 11. *Scatterplot of acidity vs size.*

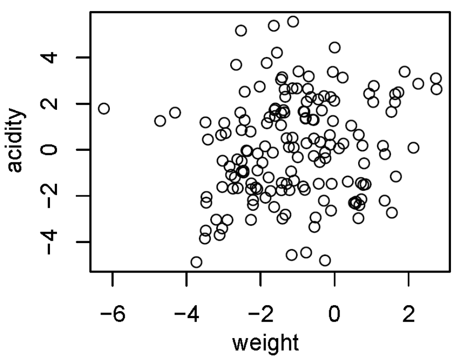

Figure 12. *Scatterplot of acidity vs weight.*

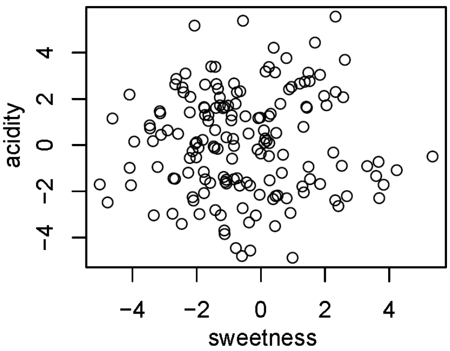

Figure 13. *Scatterplot of acidity vs sweetness.*

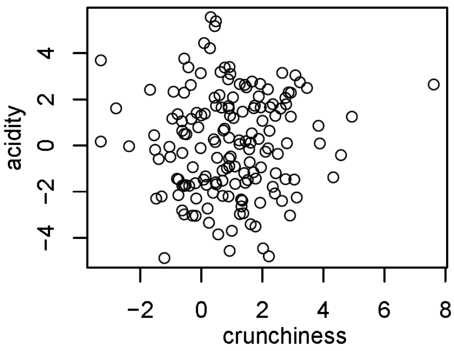

Figure 14. *Scatterplot of acidity vs crunchiness.*

Table 1. *Table showing p-values of $T_n^{(r)}$ for $r = 2, 3, 4, 5, 10$.*

B	p-values of test statistics $T_n^{(r)}$'s				
	$r=2$	$r=3$	$r=4$	$r=5$	$r=10$
100	0.0925	0.0845	0.0745	0.0715	0.0556
400	0.0635	0.0536	0.0544	0.0462	0.0448
1000	0.0578	0.0525	0.0541	0.0493	0.0388

As in *Section 5* where we observed a better power performance of $T_n^{(r)}$ with increasing r, it indicates that H_0 loses significance in that process. Hence, rejection of H_0 becomes easier for $r \uparrow$ and the concerned p-value of $T_n^{(r)}$ indeed decreases. In the above table, the p-values decrease with the increasing values of r from 2 to 10.

7. Conclusion

Under the setup of development of consistent as well as powerful robust tests for checking independence between regressors and error, the sole measure of association underlying to construction of test statistics is Kendall's τ. A test statistic based on τ is always nondegenerate and formed upon two randomly selected bivariate observations obtained on two jointly distributed random variables. However, many other measures of association had been proposed in recent past which consider more than two bivariate observations ($e.g.$ Spearman's ρ_s, *Bergsma et al.* (2014) [3]'s τ^* etc.). Also, *Bergsma* (2006) [2]'s distance based measures of association κ and ρ^*, unlike Kendall's τ can be suitable choices in this proposed setup of testing of hypothesis setup. In addition, the test statistics based on τ^* or κ are degenerate in nature, but it is possible to formulate robust test procedures on the basis of these measures by proceeding in same manner as deduced in this context. Since Kendall's τ is one of the basic nonparametric measures of association, its utility in the layout of semiparametric regression, more specifically generalized partially linear regression, has been investigated throughout this article.

Various examples on partially linear models with specified conditional error structures reveal on enhancement of the asymptotic powers of $T_n^{(r)}$ due to improved values of r. All the power curves are consistent also, as they close to 1 for $\mu \uparrow$. So, the association between jointly distributed covariates and random error is quite sensitive; falsity of the null hypothesis is well captured by $T_n^{(r)}$ as r increases. Hence, the robustness of a nonparametric statistic is important to detect the presence dependence in a partially linear model.

In real data analysis, the p-values decrease with increase in number of resampling alongside increasing order of difference r. The programs, tables, diagrams, dataset etc., are available in https://github.com/sthdas999/Asymptotic_power_performance _of_test_statistic_based_on_Kendall_s_-

Acknowledgments

The author is grateful to Dr. Saran Ishika Maiti from Visvabharati University, West Bengal, India and Dr. Ayan Pal from University of Burdwan, West Bengal, India for their valuable opinions to embellish this article. Also, the author is grateful to the anonymous referee as well as the editor of the journal for considering the article. Also, the referee's insightful advices for further betterment of this article are praiseworthy.

Conflicts of Interest

The author declares no conflicts of interest.

Appendix

Appendix I: Asymptotic Powers of $T_n^{(r)}$'s in Examples for Different Values of μ

Table 2. *Asymptotic powers of $T_n^{(r)}$'s in Example 5.1 and Example 5.2 for different values of μ where* $\underset{\sim 1}{c} = (-1.5, -1.7, 1.2, 1.3)'$.

μ	r = 2	r = 3	r = 4	r = 5	r = 10
0	0.05	0.05	0.05	0.05	0.05
5	0.1289	0.1400	0.1421	0.1682	0.2321
10	0.2682	0.3029	0.3097	0.3906	0.5719
15	0.4582	0.5192	0.5308	0.6575	0.8631
20	0.6585	0.7299	0.7426	0.8621	0.9776
25	0.8217	0.8804	0.8898	0.9619	0.9983
30	0.9244	0.9592	0.9641	0.9930	0.9999

μ	r = 2	r = 3	r = 4	r = 5	r = 10
0	0.05	0.05	0.05	0.05	0.05
5	0.1148	0.1338	0.1609	0.2152	0.3282
10	0.2244	0.2837	0.3681	0.5270	0.7749
15	0.3768	0.4858	0.6243	0.8223	0.9747
20	0.5517	0.6918	0.8342	0.9625	0.9992
25	0.7169	0.8502	0.9479	0.9958	1
30	0.8455	0.9422	0.9887	0.9998	1

Table 3. *Asymptotic powers of $T_n^{(r)}$'s in Example 5.3 and Example 5.4 for different values of μ where* $\underset{\sim 2}{c} = (-1.5, -1.7, 1.2, 1.3, -3.5)'$.

μ	r = 2	r = 3	r = 4	r = 5	r = 10
0	0.05	0.05	0.05	0.05	0.05
5	0.1403	0.1599	0.1814	0.2019	0.2852
10	0.3039	0.3651	0.4306	0.4901	0.6952
15	0.521	0.6199	0.7123	0.7838	0.9438
20	0.7319	0.8303	0.9023	0.9447	0.9962
25	0.8818	0.9458	0.9788	0.9919	0.9999
30	0.9599	0.9879	0.9972	0.9993	1

μ	r = 2	r = 3	r = 4	r = 5	r = 10
0	0.05	0.05	0.05	0.05	0.05
5	0.1198	0.1479	0.1682	0.1774	0.2553
10	0.2397	0.3279	0.3906	0.4185	0.6291
15	0.4057	0.5611	0.6575	0.6963	0.906
20	0.5911	0.7743	0.8621	0.8913	0.9894
25	0.7577	0.9119	0.9619	0.9746	0.9995
30	0.8786	0.9745	0.9930	0.9962	1

Table 4. *Asymptotic powers of $T_n^{(r)}$'s in Example 5.5 and Example 5.6 for different values of μ where* $\underset{\sim 3}{c} = (2.5, 4, -6.2, -5.5, 3.1)'$.

μ	r = 2	r = 3	r = 4	r = 5	r = 10
0	0.05	0.05	0.05	0.05	0.05
5	0.1368	0.1706	0.2189	0.3102	0.3287
10	0.2929	0.3979	0.5371	0.7436	0.7758
15	0.5021	0.6681	0.8319	0.9644	0.9750
20	0.7105	0.8702	0.9664	0.9984	0.9992
25	0.8654	0.9657	0.9965	1	1
30	0.9511	0.9941	0.9998	1	1

μ	r = 2	r = 3	r = 4	r = 5	r = 10
0	0.05	0.05	0.05	0.05	0.05
5	0.1074	0.1577	0.1838	0.1996	0.2058
10	0.2016	0.3582	0.4376	0.4835	0.5010
15	0.3331	0.6094	0.7214	0.7764	0.7957
20	0.4892	0.8208	0.9084	0.9408	0.9505
25	0.6471	0.9405	0.9810	0.9911	0.9933
30	0.7828	0.9861	0.9976	0.9992	0.9995

Table 5. *Asymptotic powers of $T_n^{(r)}$'s in Example 5.7 and Example 5.8 for different values of μ where* $\underset{\sim 4}{c} = (5.92, -3.78, -10.66, 8.89, -5.45, 9.65, 8.35, -7.89)'$.

μ	r = 2	r = 3	r = 4	r = 5	r = 10
0	0.05	0.05	0.05	0.05	0.05
5	0.1681	0.2435	0.2685	0.4218	0.611
10	0.3904	0.6005	0.6591	0.8944	0.9864
15	0.6573	0.8858	0.9247	0.9965	1
20	0.8618	0.9844	0.9931	1	1
25	0.9618	0.9990	0.9998	1	1
30	0.9932	1	1	1	1

μ	r = 2	r = 3	r = 4	r = 5	r = 10
0	0.05	0.05	0.05	0.05	0.05
5	0.1091	0.1517	0.1739	0.2916	0.3258
10	0.2064	0.3396	0.4078	0.7081	0.7708
15	0.3423	0.5800	0.6818	0.9499	0.9735
20	0.5026	0.7932	0.8807	0.9969	0.9991
25	0.6625	0.9241	0.9702	0.9999	1
30	0.7974	0.9798	0.9952	1	1

Appendix II: Combined Proofs of Proposition 3.1 and Theorem 3.1

Note that, for $i = 1, \ldots, n$ and $\delta > 0$,

$$\begin{aligned} P\left(|(Y_{i+1} - Y_i) - (\epsilon_{i+1} - \epsilon_i)| > \delta\right) &= P\left(|\underset{\sim}{\beta}^T(\underset{\sim i+1}{X} - \underset{\sim i}{X}) + \{m(\underset{\sim i+1}{W}) - m(\underset{\sim i}{W})\}| > \delta\right) \\ &\leq P\left(|\sum_{s=1}^{p} \beta_s (X_{i+1\,s} - X_{i\,s})| + |m(\underset{\sim i+1}{W}) - m(\underset{\sim i}{W})| > \delta\right). \end{aligned} \tag{16}$$

Since $m(\cdot, \ldots, \cdot)$ is Lipschitz continuous on $\mathbb{R}^q$, therefore $|m(\underset{\sim i+1}{W}) - m(\underset{\sim i}{W})| \leq C \cdot ||\underset{\sim i+1}{W} - \underset{\sim i}{W}||_q$, $C > 0$.

$\therefore$

(16) $\leq P\left(|\sum_{s=1}^{p} \beta_s (X_{i+1\,s} - X_{i\,s})| + C \cdot ||\underset{\sim i+1}{W} - \underset{\sim i}{W}||_q > \delta\right) = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} P\left(||\underset{\sim i+1}{W} - \underset{\sim i}{W}||_q > \delta^{'}\right) \times \prod_{s=1}^{p} dH_{D_{is}}(d_{is})$,

where $D_{is} = X_{i+1\,s} - X_{i\,s}$ for $s = 1, \ldots, p$ and $\delta^{'} = C^{-1}\left(\delta - |\sum_{s=1}^{p} \beta_s d_{is}|\right)$. Moreover,

$$P\left[\left(\sum_{m=1}^{q} |W_{i+1\,m} - W_{i\,m}|^q\right)^{\frac{1}{q}} > \delta^{'}\right] = P\left[|W_{i+1\,1} - W_{i\,1}| > \left(\delta^{'q} - \left(\sum_{m=2}^{q} |W_{i+1\,m} - W_{i\,m}|^q\right)\right)^{\frac{1}{q}}\right] \tag{17}$$

for $m = 2, \ldots, q;\ i = 1, \ldots, n-1$. Next, the model is re-expressed by ordering the W_1-observations as

$$Y_K^* = \beta_1 X_{K1}^* + \ldots + \beta_p X_{Kp}^* + m(W_{K1}^*, \ldots, W_{Kq}^*) + \epsilon_K^*;\ \ K = 1, \ldots, n$$

where $\{W_{11}^*, \ldots, W_{n1}^*\}$ are the n observations on W_1 such that $W_{11}^* \leq \ldots \leq W_{n1}^*$. Corresponding to the ordered observations of W_1, the observations on $W_2, \ldots, W_q$ as well as $X_1, \ldots, X_p$ are the induced ordered observations. The responses $\{Y_1^*, \ldots, Y_n^*\}$ are called induced ordered responses corresponding to $\{W_{11}^*, \ldots, W_{n1}^*\}$.

Then, (17) is further deduced as $\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} P\left[|W_{i+1\,1}^* - W_{i\,1}^*| > \delta^{''}\right] \times \prod_{m=2}^{q} dF_{T_{im}^*}(t_{im}^*)$, where $T_{im}^* = W_{i+1\,m}^* - W_{i\,m}^*$ and $\delta^{''} = \left(\delta^{'q} - \left(\sum_{m=2}^{q} |t_{im}^*|^q\right)\right)^{\frac{1}{q}}$. Observe that, $(W_{i+1\,1}^* - W_{i\,1}^*)$ is the i-th spacing (Pyke (1965))[**?**] on W_1, $i = 1, \ldots, n-1$. Suppose $F_{W_1}(\cdot)$ is the CDF of W_1. For any $\delta^* > 0$, it is possible to deduce $P\left[|W_{i+1\,1}^* - W_{i\,1}^*| > \delta^*\right]$ as $P\left[\frac{n\Delta_{n-1;n}}{\log\, n} > \frac{nR}{\log\, n}\right]$, where $\Delta_{n-1;n}$ is the maximal spacing based on $(n-1)$ uniform spacings $(U_{i+1\,1}^* - U_{i1}^*)$s and $R = \inf_{1 \leq i \leq n-1} \delta^*/|F_{W_1}^{-1'}(\xi_{i;i+1})|$. Due to to *Lévy* (1939) [13], one can verify that

$$P\left[\frac{n\Delta_{n-1;n}}{\log\, n} > \frac{nR}{\log\, n}\right] \longrightarrow 1 - \exp\left(-\exp\left(-\frac{nR}{\log\, n}\right)\right). \tag{18}$$

Therefore, $\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} P\left[|W_{i+1\,1}^* - W_{i\,1}^*| > \delta^{''}\right] \times \prod_{m=2}^{q} dF_{T_{im}^*}(t_{im}^*) \leq 1 - \exp\left(-\exp\left(-\frac{nR}{\log\, n}\right)\right) \longrightarrow 0$, which further implies the R.H.S. of (18) tends to $1 - e^0 = 0$. Then, $|W_{i+1\,1}^* - W_{i\,1}^*| = o_p(1) \implies \sup_{i \in \{1, \ldots, n-1\}} |W_{i+1\,1}^* - W_{i\,1}^*| = o_p(1)$.

Hence, $P\left(|(Y_{i+1}^* - Y_i^*) - (\epsilon_{i+1}^* - \epsilon_i^*)| > \delta\right) \longrightarrow 0$, $i = 1, \ldots, n-1$, where $(Y_{i+1}^* - Y_i^*)$ is the first order difference of Y^* and $(\epsilon_{i+1}^* - \epsilon_i^*)$ the first order difference of ϵ^*. Next, we need to verify if $\epsilon^*(r) \approx Y^*(r)$, where $Y^*(r)$ is the r-th order difference of Y and $\epsilon^*(r)$ is the r-th order difference of ϵ. Define $L(\epsilon_1, \ldots, \epsilon_{r+1}) = \alpha_1 \epsilon_1 + \ldots + \alpha_{r+1} \epsilon_{r+1}$ with $\alpha_1, \ldots, \alpha_{r+1} \in \mathbb{Z}$ and $\sum_{i=1}^{r+1} \alpha_i = 0$, $\epsilon_1, \ldots, \epsilon_{r+1}$ are $(r+1)$ $i.i.d.$ errors.

The k-th order absolute raw moment of $L(\epsilon_1, \ldots, \epsilon_{r+1})$ is

$$E\Big|L(\epsilon_1, \ldots, \epsilon_{r+1})\Big|^k = E\Big|\sum_{i=1}^{r+1} \alpha_i \epsilon_i\Big|^k \le E\left[\sqrt{\sum_{i=1}^{r+1} \alpha_i^2} \cdot \sqrt{\sum_{i=1}^{r+1} \epsilon_i^2}\right]^k \text{ (using Cauchy-Schwartz inequality)}$$

$$= \left(\sum_{i=1}^{r+1} \alpha_i^2\right)^{k/2} E\left[\sqrt{\sum_{i=1}^{r+1} \epsilon_i^2}\right]^k = \eta(\alpha_1, \ldots, \alpha_{r+1}) \cdot E\left[\sqrt{\sum_{i=1}^{r+1} \epsilon_i^2}\right]^k \qquad (19)$$

where $\eta(\alpha_1, \ldots, \alpha_{r+1}) = \left(\sum_{i=1}^{r+1} \alpha_i^2\right)^{k/2}$. Observe that, $\eta(\alpha_1, \ldots, \alpha_{r+1}) \le \max_{\alpha_1, \ldots, \alpha_{r+1} \neq 0} \eta(\alpha_1, \ldots, \alpha_{r+1})$ with $\sum_{i=1}^{r+1} \alpha_i = 0$. Taking $r = 2$, one can deduce $\eta(\alpha_1, \alpha_2, \alpha_3) = 2\left(\dfrac{\alpha_3^3 - \alpha_1^3}{\alpha_3 - \alpha_1}\right) = S(\alpha_1, \alpha_3)$, say. Then, $\log S = \log 2 + \log(\alpha_3^3 - \alpha_1^3) - \log(\alpha_3 - \alpha_1)$. Maximizing $\log S$ with respect to α_1, α_3 *i.e.* solving the following equations

$$\frac{\partial \log S}{\partial \alpha_1} = 0 \implies -2\alpha_1^2 + \alpha_1\alpha_3 + \alpha_3^2 = 0 \qquad (20)$$

$$\text{and} \quad \frac{\partial \log S}{\partial \alpha_3} = 0 \implies 2\alpha_3^2 - \alpha_1^2 - \alpha_1\alpha_3 = 0 \qquad (21)$$

one gets $\alpha_1 = \pm\alpha_3$. But α_i's $\neq 0$, hence $\alpha_1 = \alpha_3 \implies \alpha_2 = -2\alpha_1$. So, $\eta(\alpha_1, \alpha_2, \alpha_3)$ has maximum value at $(\alpha_1, -2\alpha_1, \alpha_1)$. Also, from (19), the maximum value of k-th order absolute moment of $(\alpha_1\epsilon_1 - 2\alpha_1\epsilon_2 + \alpha_1\epsilon_3)$ satisfies $E|\alpha_1\epsilon_1 - 2\alpha_1\epsilon_2 + \alpha_1\epsilon_3|^k \le |\alpha_1|^k\{1^2 + (-2)^2 + 1^2\}^{k/2} E\left(\sqrt{\sum_{i=1}^{3} \epsilon_i^2}\right)^k$, for any nonzero integer α_1. If $\alpha_1 = \pm 1$, the linear contrast $\pm(\epsilon_1 - 2\epsilon_2 + \epsilon_3)$ has minimum variance as well as maximum k-th absolute raw moment among all possible linear contrasts $(\delta_1\epsilon_1 + \delta_2\epsilon_2 + \delta_3\epsilon_3)$. Here $L(\epsilon_1, \epsilon_2, \epsilon_3) = \epsilon_1 - 2\epsilon_2 + \epsilon_3$ is termed as the *second order difference of* ϵ based on $\epsilon_1, \epsilon_2, \epsilon_3$, which is indeed a first order difference of two first order differences of ϵ, i.e. $(\epsilon_1 - \epsilon_2) - (\epsilon_2 - \epsilon_3) = \epsilon_1^*(1) - \epsilon_2^*(1) \approx Y_1^*(1) - Y_2^*(1), \implies \epsilon^*(2) \approx Y^*(2)$.

Similarly, one can verify that for $r = 3$, $\eta(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_4^2$ subjected to $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 0$ has maximum value at $(-\alpha_4, 3\alpha_4, \alpha_4)$, $\alpha_4 \neq 0$ and $E| - 3\alpha_4\epsilon_1 - \alpha_4\epsilon_2 + 3\alpha_4\epsilon_3 + \alpha_4\epsilon_4|^k \ge E|\lambda_1\epsilon_1 + \lambda_2\epsilon_2 + \lambda_3\epsilon_3 + \lambda_4\epsilon_4|^k$; $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}$, i.e. $|\alpha_4|^k E|\epsilon_4 - 3\epsilon_1 + 3\epsilon_3 - \epsilon_2|^k$ is maximum among the k-th order absolute moment of all possible linear functions $(\lambda_1\epsilon_1 + \lambda_2\epsilon_2 + \lambda_3\epsilon_3 + \lambda_4\epsilon_4)$. For $\alpha_4 = \pm 1$, the linear contrast $\pm(\epsilon_4 - 3\epsilon_1 + 3\epsilon_4 - \epsilon_2)$ [or equivalently, $\pm(\epsilon_1 - 3\epsilon_2 + 3\epsilon_3 - \epsilon_4)$ as the errors are i.i.d.] has the minimum variance as well as maximum k-th order moment among all the k-th order absolute moments of linear functions $(\lambda_1\epsilon_1 + \lambda_2\epsilon_2 + \lambda_3\epsilon_3 + \lambda_4\epsilon_4)$. The function $(\epsilon_1 - 3\epsilon_2 + 3\epsilon_3 - \epsilon_4)$ is denoted as the *third order difference of* ϵ based on four i.i.d. observations ϵ_1, ϵ_2, ϵ_3 and ϵ_4, which is the first order difference of two second order differences of ϵ as $\{(\epsilon_1 - 2\epsilon_2 + \epsilon_3) - (\epsilon_2 - 2\epsilon_3 + \epsilon_4)\}$, or the second order difference of three first order differences of ϵ's as $\{(\epsilon_1 - \epsilon_2) - 2(\epsilon_2 - \epsilon_3) + (\epsilon_3 - \epsilon_4)\}$. Similarly, $\epsilon^*(3) \approx Y^*(3)$.

Therefore, the second and third order differences of i.i.d. errors constitute best possible linear functions having highest second and third order absolute moments among all possible linear functions of errors respectively. It can be concluded finally that for a general order difference r of ϵ, $\epsilon^*(r) \approx Y^*(r)$. Also, $\epsilon^*(r)$ has the maximal k-th order absolute moment among all possible linear functions of $\epsilon_1, \ldots, \epsilon_{r+1}$.

Proof of Proposition 3.2

It is to be noted that for $p = 1 = q$, $\widehat{\beta} = \left(\sum_{i=1}^{n} \widehat{\epsilon}_{Xi}\widehat{\epsilon}_{Xi}^T\right)^{-1}\left(\sum_{i=1}^{n} \widehat{\epsilon}_{Xi}\widehat{\epsilon}_{Yi}\right)$ where $\widehat{\epsilon}_{Yi} = Y_i - \widehat{m}_Y(W_i)$ and $\widehat{\epsilon}_{Xi} = X_i - \widehat{m}_X(W_i)$ based on random sample of size n (Y_i, X_i, W_i), $i = 1, \ldots, n$, from (Y, X, W). Then,

$$\widehat{Y} = X\,\widehat{\beta} + \widehat{m}\,(W) = S(Y, X, W) = S(X\beta + m(W) + \epsilon, X, W) = S(Z + \epsilon, X, W), \text{ say, where } Z = X\beta + m(W).$$

Using Taylor's theorem, the expansion of S upto first order approximation is given by

$$\begin{aligned} S(X\beta + m(W) + \epsilon, X, W) &\simeq S(X\beta + m(W), X, W) + \begin{pmatrix} \epsilon \\ 0 \\ 0 \end{pmatrix}^T \begin{pmatrix} \frac{\partial S}{\partial Z} \\ \frac{\partial S}{\partial X} \\ \frac{\partial S}{\partial W} \end{pmatrix}\Bigg|_{\substack{Z=\epsilon \\ X=0 \\ W=0}} \\ &= S(X\beta + m(W), X, W) + \epsilon \cdot \left(\frac{\partial S}{\partial Z}\Big|_{Z=\epsilon} \right). \end{aligned}$$

Assume that $\sup\limits_{X,W \in \mathbb{R}} \left| \frac{\partial}{\partial Z} S(Z, X, W)\Big|_{Z=\epsilon} \right| < \infty$, say $\mathbf{A}$. It is to be noted that

$$S^{(r)}(X\beta + m(W) + \epsilon, X, W) \simeq S^{(r)}(X\beta + m(W), X, W) + \left[\epsilon \cdot \left(\frac{\partial S}{\partial Z}\Big|_{Z=\epsilon} \right) \right](r)$$

where $S^{(r)}$ denotes the r-th order difference of S and $\left[\epsilon \cdot \left(\frac{\partial S}{\partial Z}\Big|_{Z=\epsilon} \right) \right](r)$ is the r-th order difference of $\epsilon \cdot \left(\frac{\partial S}{\partial Z}\Big|_{Z=\epsilon} \right)$.

Furthermore, $\sup\limits_{X,W,\epsilon \in \mathbb{R}} \left| \epsilon \cdot \left(\frac{\partial S}{\partial Z}\Big|_{Z=\epsilon} \right) \right| = \mathbf{A}\epsilon$ that implies $\left| \left[\epsilon \cdot \left(\frac{\partial S}{\partial Z}\Big|_{Z=\epsilon} \right) \right](r) \right| = \mathbf{A}\epsilon^*(r)$.

Then,

$$S^{(r)}(X\beta + m(W) + \epsilon, X, W) - S^{(r)}(X\beta + m(W), X, W) \simeq \left[\epsilon \cdot \left(\frac{\partial S}{\partial Z}\Big|_{Z=\epsilon} \right) \right](r)$$

$$\implies P\left(|S^{(r)}(X\beta + m(W) + \epsilon, X, W) - S^{(r)}(X\beta + m(W), X, W)| > \delta \right) \simeq P\left(\left| \left[\epsilon \cdot \left(\frac{\partial S}{\partial Z}\Big|_{Z=\epsilon} \right) \right](r) \right| > \delta \right).$$

Now, it is worth to realize that

$$P\left(\left| \left[\epsilon \cdot \left(\frac{\partial S}{\partial Z}\Big|_{Z=\epsilon} \right) \right](r) \right| > \delta \right) \le P\left(\left| \mathbf{A}\epsilon^*(r) \right| > \delta \right) = P\left(\left| \epsilon^*(r) \right| > \delta/|\mathbf{A}| \right) = P\left(\epsilon^*(r) > \delta/|\mathbf{A}| \right) + P\left(\epsilon^*(r) < -\delta/|\mathbf{A}| \right).$$

Using Markov's inequality, we get $P\left(\epsilon^*(r) > \delta/|\mathbf{A}| \right) \le \dfrac{E[\epsilon^*(r)]}{\delta/|\mathbf{A}|} = \dfrac{|\mathbf{A}|}{\delta} \sum\limits_{j=1}^{r+1} (-1)^{j-1} \binom{r}{j-1} E(\epsilon_j)$,
where $E(\epsilon_j) = E_{X,W} E(\epsilon_j | X, W) = 0$. Then, $P\left(\epsilon^*(r) > \delta/|\mathbf{A}| \right) = 0$. In similar manner, $P\left(\epsilon^*(r) < -\delta/|\mathbf{A}| \right) = 0$.
$\therefore P\left(|S^{(r)}(X\beta + m(W) + \epsilon, X, W) - S^{(r)}(X\beta + m(W), X, W)| > \delta \right) \simeq 0 \implies S^{(r)}(X\beta + m(W) + \epsilon, X, W) \approx S^{(r)}(X\beta + m(W), X, W)$. The proposition can be proved in similar way for $p, q > 1$.

Proof of Theorem 4.1 The variance of kernel of $T_n^{(r)}$ is computed under H_0 as

$$\begin{aligned} \eta_{1,2}(r) &= Var\left[E(h((\widehat{Y}_1^*(r), Y_1^*(r)), (\widehat{Y}_2^*(r), Y_2^*(r))) \Big| (\widehat{Y}_1^*(r), Y_1^*(r))) \right] = Var\left[\psi(\widehat{Y}_1^*(r), Y_1^*(r)) \right], \text{ say, where} \\ \psi(\widehat{Y}_1^*(r), Y_1^*(r)) &= E\left[h((\widehat{Y}_1^*(r), Y_1^*(r)), (\widehat{Y}_2^*(r), Y_2^*(r))) \Big| (\widehat{Y}_1^*(r), Y_1^*(r)) \right] \\ &= P\left[\widehat{Y}_1^*(r) > Y_1^*(r) > \widehat{Y}_2^*(r), Y_2^*(r) | \widehat{Y}_1^*(r), Y_1^*(r) \right] + P\left[\widehat{Y}_1^*(r) < Y_1^*(r) < \widehat{Y}_2^*(r), Y_2^*(r) | \widehat{Y}_1^*(r), Y_1^*(r) \right] \\ &\quad - P\left[\widehat{Y}_1^*(r) > Y_1^*(r) < \widehat{Y}_2^*(r), Y_2^*(r) | \widehat{Y}_1^*(r), Y_1^*(r) \right] - P\left[\widehat{Y}_1^*(r) < Y_1^*(r) > \widehat{Y}_2^*(r), Y_2^*(r) | \widehat{Y}_1^*(r), Y_1^*(r) \right] \end{aligned}$$

which is finally deduced as $\left(2F_{0;\widehat{Y}^*(r)}(\widehat{Y}_1^*(r)) - 1 \right) \left(2F_{0;Y^*(r)}(Y_1^*(r)) - 1 \right)$. $F_{0;\widehat{Y}^*(r)}(\cdot)$ and $F_{0;Y^*(r)}(\cdot)$ are the marginal CDFs of $\widehat{Y}^*(r)$ and $Y^*(r)$ respectively under H_0. Then,

$$\begin{aligned} Var\left[\psi(\widehat{Y}_1^*(r), Y_1^*(r)) \right] &= Var\left[\left(2F_{0;\widehat{Y}^*(r)}(\widehat{Y}_1^*(r)) - 1 \right) \left(2F_{0;Y^*(r)}(Y_1^*(r)) - 1 \right) \right] \\ &= E\left(2F_{0;\widehat{Y}^*(r)}(\widehat{Y}_1^*(r)) - 1 \right)^2 E\left(2F_{0;Y^*(r)}(Y_1^*(r)) - 1 \right)^2 \\ &\quad - E^2\left(2F_{0;\widehat{Y}^*(r)}(\widehat{Y}_1^*(r)) - 1 \right) E^2\left(2F_{0;Y^*(r)}(Y_1^*(r)) - 1 \right) \end{aligned}$$

Since $F_{0;\widehat{Y}^*(r)}(\cdot)$, $F_{0;Y^*(r)}(\cdot) \overset{indep.}{\sim} U(0,1)$, therefore $E\left(2F_{0;\widehat{Y}^*(r)}(\widehat{Y}_1^*(r)) - 1\right)^2 = 4 \cdot \frac{1}{12} = \frac{1}{3} = E\left(2F_{0;Y^*(r)}(Y_1^*(r)) - 1\right)^2$.

Also, $E\left(2F_{0;\widehat{Y}^*(r)}(\widehat{Y}_1^*(r)) - 1\right) = E\left(2F_{0;Y^*(r)}(Y_1^*(r)) - 1\right) = 0$. Therefore, $\eta_{1,2}(r) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$.

Proof of Proposition 5.1

For sample size n^* such that $n^* > n$,

$$\begin{aligned} P_{H_n}\left[\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) > t_\kappa\right] &= P_{H_n}\left[\sqrt{\frac{n^*}{n}} \cdot \sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) > \sqrt{\frac{n^*}{n}} \cdot t_\kappa\right] \\ &= P_{H_n}\left[\sqrt{n^*}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) > \sqrt{\frac{n^*}{n}} \cdot t_\kappa\right] \\ &< P_{H_n}\left[\sqrt{n^*}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) > t_\kappa\right] \overset{asy.}{\equiv} P_{H_n}\left[\sqrt{n^*}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) > t_\kappa\right], \end{aligned}$$

$i.e.$ for increasing sample size n, the power of $T_n^{(r)} \uparrow$ and tends to 1. Moreover, as $\mu \uparrow$,

$$P_{H_n}\left[\sqrt{n}(T_n^{(r)} - E_{H_0}(T_n^{(r)})) > t_\kappa\right] = \Phi\left(\frac{\mu\, E_{H_1}(T_n^{(r)}) - t_\kappa}{\sqrt{4\eta_{1,2}(r)}}\right) \longrightarrow 1.$$

Table 6. *Table showing various order differences of some values.*

Values	0.119, -0.8029, 0.016, 0.11, 0.922, 0.494, 1.101, -1.210, -0.133, -0.257
First order differences	-0.921, 0.818, 0.097, 0.809, -0.429, 0.607, -2.311, 1.077, -0.123
Second order differences	-0.103, 0.916, 0.907, 0.380, 0.179, -1.704, -1.234, 0.954
Third order differences	-0.005, 1.725, 0.478, 0.988, -2.133, -0.627, -1.358
Fourth order differences	0.804, 1.296, 1.085, -1.324, -1.056, -0.750
Fifth order differences	0.375, 1.903, -1.226, -0.247, -1.179

References

[1] Andrews, D. W. (1995). Nonparametric kernel estimation for semiparametric models. Econometric Theory, 11(3): 560-586. https://doi.org/10.1017/S0266466600009427

[2] Bergsma, W. P. (2006). A new correlation coefficient, its orthogonal decomposition and associated tests of independence. https://doi.org/10.48550/arXiv.math/0604627

[3] Bergsma, W. and Dassios, A. (2014). A consistent test of independence based on a sign covariance related to Kendall's tau. Bernoulli, 1006-1028. DOI: 10.3150/13-BEJ514

[4] Cam Le, L. (1960). Locally asymptotically normal families of distributions. University of California Publications in Statistics, 3: 37-98. https://doi.org/10.1007/978-1-4684-0377-0_5

[5] Cuzick, J. (2014). Semiparametric additive regression. Journal of the Royal Statistical Society: Series B (Methodological), 54(3): 831-843. https://doi.org/10.1111/j.2517-6161.1992.tb01455.x

[6] Das, S. & Maiti, S. I. (2022). On the Test of Association Between Nonparametric Covariate and Error in Semiparametric Regression Model. Journal of the Indian Society for Probability and Statistics, 23(2): 541-564. https://doi.org/10.1007/s41096-022-00139-0

[7] Das, S., Halder, S., and Maiti, S. I. (2023). An Extended Approach to Test of Independence between Error and Covariates under Nonparametric Regression Model. Thailand Statistician, 21(1): 19-36.

[8] Dhar, S. S., Dassios, A., and Bergsma, W. (2018). Testing Independence of Covariates and Errors in Nonparametric regression. Scandinavian Journal of Statistics, 45: 421-443.

[9] Einmahl JH, Van Keilegom I. (2008) Specification tests in nonparametric regression. Journal of Econometrics, 143(1): 88-102

[10] Hamilton, S. A., & Truong, Y. K. (1997). Local linear estimation in partly linear models. Journal of Multivariate Analysis, 60(1): 1-19. https://doi.org/10.1006/jmva.1996.1642

[11] Kendall, M. G. (1938). A new measure of rank correlation. Biometrika, 30(1/2): 81-93. https://doi.org/10.1093/biomet/30.1-2.81

[12] Liu, Z., Liu, Z., Lu, X., & Lu, X. (1997). Root-n-consistent semiparametric estimation of partially linear models based on k-nn method. Econometric Reviews, 16(4): 411-420. https://doi.org/10.1080/07474939708800396

[13] Lévy, P. (1939). Sur la division d'un segment par des points choisis au hasard. CR Acad. Sci. Paris, 208: 147-149.

[14] Li, Q. (2000). Efficient estimation of additive partially linear models. International Economic Review, 41(4): 1073-1092. https://doi.org/10.1111/1468-2354.00096

[15] Pyke, R. (1965). Spacings. Journal of the Royal Statistical Society: Series B (Methodological), 27(3): 395-436. https://doi.org/10.1111/j.2517-6161.1965.tb00602.x

[16] Robinson, P. M. (1988). Root-N-consistent semiparametric regression. Econometrica, 56(4): 931-954. https://doi.org/10.2307/1912705

[17] Silverman, Bernard W. (2018). Density estimation for statistics and data analysis. Routledge. https://doi.org/10.1201/9781315140919

[18] Szekely, G. J., Rizzo, M. L., and Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. The Annals of Statistics, 35(6): 2769-2794. https://doi.org/10.1214/009053607000000505

[19] Wang, L., Brown, L. D., & Cai, T. T. (2011). A difference based approach to the semiparametric partial linear model. Electron. J. Statist., 5: 619-641. https://doi.org/10.1214/11-EJS621

[20] Zhou, Z., Mentch, L., & Hooker, G. (2021). V-statistics and variance estimation. Journal of Machine Learning Research, 22(287): 1-48.