Testing Equality of Two Intercepts for the Parallel Regression Model with Non-sample Prior Information

Budi Pratikno[1] and Shahjahan Khan[2]
[1] Department of Mathematics and Natural Science
Jenderal Soedirman University, Purwokerto, Jawa Tengah, Indonesia
[2] School of Agricultural, Computational and Environmental Sciences
Centre for Sustainable Catchments, University of Southern Queensland
Toowoomba, Queensland, Australia
Email: *b_pratikto@yahoo.com.au* and *khans@usq.edu.au*

Abstract

This paper proposes tests for equality of intercepts of two simple regression models when non-sample prior information (NSPI) is available on the equality of two slopes. For three different scenarios on the values of the slope, namely (i) unknown (unspecified), (ii) known (specified), and (iii) suspected, we derive the unrestricted test (UT), restricted test (RT) and pre-test test (PTT) for testing equality of intercepts. The test statistics, their sampling distributions, and power functions of the tests are obtained. Comparison of power function and size of the tests reveal that the PTT has a reasonable dominance over the UT and RT.

Keywords and phrases: Linear regression; intercept and slope parameters; pre-test test; non-sample prior information; and power function.

2010 Mathematics Subject Classification: Primary 62F03 and Secondary 62J05.

1 Introduction

Inferences about population parameters could be improved using non-sample prior information (NSPI) from trusted sources (cf Bancroft, 1944). Such information are usually available from previous studies or expert knowledge or experience of the researchers, and are unrelated to any sample data.

It is well known that, for any linear regression model, the inference on the intercept parameter depends on the value of the slope parameter. Thus the non-sample prior information on the value of the slope parameter would directly affect the inference on the intercept parameter.

An appropriate statistical test on the suspected value of the slopes, after expressing it in the form a null hypothesis, is useful to eliminate the uncertainty on this suspected information. Then the outcome of the preliminary test on the uncertain NSPI on the slopes is used in the hypothesis testing on the intercepts to improve the performance of the statistical test (cf. Khan and Saleh, 2001; Saleh, 2006, p. 55-58; Yunus and Khan, 2011a).

As an example, in any *spotlight analysis* the aim is to compare the mean responses of the two categorical groups at specific values of the continuous covariate. Furthermore, we consider a response variable (η), a continuous covariate (χ) and a categorical explanatory variable (ζ) with two categories (eg treatment and control). If there is an association between χ and ζ, the least squares line of η on χ will be parallel with different intercepts for two different categories of ζ. However, the two fitted lines will not be parallel if there is no association between the two explanatory variables because of the presence of interaction. The scenario will be different if the two explanatory variables are associated and they also interact.

In any inference, estimation or test, on the equality of the two intercepts of the two regression lines of Y on X for two different categories of Z, the slope of the regression lines plays a key role. The test (also the estimation) of intercept is directly impacted by the values of the slope. Therefore, the type of NSPI on the value of the slopes will influence the inference on the intercepts.

The suspected NSPI on the slopes may be (i) unknown or unspecified if NSPI is not available, (ii) known or specified if the exact value is available from NSPI, and (iii) uncertain if the suspected value is unsure. For the three different scenarios, three different statistical tests, namely the (i) unrestricted test (UT), (ii) restricted test (RT) and (iii) pre-test test (PTT) are defined.

In the area of estimation with NSPI there has been a lot of work, notably Bancroft (1944, 1964), Hand and Bancroft (1968), and Judge and Bock (1978) introduced a preliminary test estimation of parameters to estimate the parameters of a model with uncertain prior information. Khan (2000, 2003, 2005, 2008), Khan and Saleh (1997, 2001, 2005, 2008), Khan et al. (2002), Khan and Hoque (2003), Saleh (2006) and Yunus (2010) covered various work in the area of improved estimation using NSPI, but there is a very limited number of studies on the testing of parameters in the presence of uncertain NSPI. Although Tamura (1965), Saleh and Sen (1978, 1982), Yunus and Khan (2007, 2011a, 2011b), and Yunus (2010) used the NSPI for testing hypotheses using nonparametric methods, the problem has not been addressed in the parametric context.

A parallelism problem can be described as a special case of two related regression lines on the same dependent and independent variables that come from two different categories of the respondents. If the independent data sets come from two random samples, researchers often wish to model the regression lines that are parallel (i.e. the slopes of the two regression lines are equal) or check whether the lines have the same intercept on the vertical-axis. To test the parallelism of the two regression equations, namely

$$y_{1j} = \theta_1 + \beta_1 x_{1j} + e_{1j} \text{ and } y_{2j} = \theta_2 + \beta_2 x_{2j} + e_{2j},\ j\text{=}1,2,\cdots,n_i,$$

for the two data sets: $\boldsymbol{y} = [\boldsymbol{y}_1^{'}, \boldsymbol{y}_2^{'}]^{'}$ and $\boldsymbol{x} = [\boldsymbol{x}_1^{'}, \boldsymbol{x}_2^{'}]^{'}$ where $\boldsymbol{y}_1 = [y_{11}, \cdots, y_{1n_1}]^{'}$, $\boldsymbol{y}_2 = [y_{21}, \cdots, y_{2n_2}]^{'}$, $\boldsymbol{x}_1 = [x_{11}, \cdots, x_{1n_1}]^{'}$ and $\boldsymbol{x}_2 = [x_{21}, \cdots, x_{2n_2}]^{'}$, we use an appropriate two-sample t test for testing $H_0 : \beta_1 = \beta_2$ (parallelism). This t statistic is given as

$$t = \frac{\widetilde{\beta_1} - \widetilde{\beta_2}}{S_{(\widetilde{\beta_1} - \widetilde{\beta_2})}},$$

where $\widetilde{\beta_1}$ and $\widetilde{\beta_2}$ are estimate of the slopes β_1 and β_2 respectively, and $S_{(\widetilde{\beta_1} - \widetilde{\beta_2})}$ is the standard error of the estimated difference between the two slopes (Kleinbaum et al., 2008, p. 223). The parallelism of the two regression equations above can be expressed as a single model in matrix form, that is,

$$\boldsymbol{y} = \boldsymbol{X\Phi} + \boldsymbol{e},$$

where $\boldsymbol{\Phi} = [\theta_1, \theta_2, \beta_1, \beta_2]^{'}$, $\boldsymbol{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2]^{'}$ with $\boldsymbol{X}_1 = [1, 0, x_1, 0]^{'}$ and $\boldsymbol{X}_2 = [0, 1, 0, x_2]^{'}$ and $\boldsymbol{e} = [e_1, e_2]^{'}$. The matrix form of the intercept and slope parameters can be written, respectively, as $\boldsymbol{\theta} = [\theta_1, \theta_2]^{'}$ and $\boldsymbol{\beta} = [\beta_1, \beta_2]^{'}$ (cf Khan, 2006).

For the model under study two independent bivariate samples are considered such that $y_{ij} \sim N(\theta_i + \beta_i x_{ij}, \sigma^2)$ for $i = 1, 2$ and $j = 1, \cdots, n_i$. See Khan (2003, 2006, 2008) for details on parallel regression models and related analyses.

To explain the importance of testing the equality of the intercepts when the equality of slopes is uncertain, we consider the general form of the two parallel simple regression models (PRM) as follows

$$\boldsymbol{Y}_i = \theta_i \boldsymbol{1}_{n_i} + \beta_i \boldsymbol{x}_{ij} + \boldsymbol{e}_{ij},\ i = 1, 2, \text{ and } j\text{=}1,2,\cdots,n_i, \tag{1.1}$$

where $\boldsymbol{Y}_i = (Y_{i1}, \cdots, Y_{in_i})^{'}$ is a vector of n_i observable random variables, $\boldsymbol{1}_{n_i} = (1, \cdots, 1)^{'}$ is an n_i-tuple of $1^{'}s$, $\boldsymbol{x}_{ij} = (x_{i1}, \cdots, x_{in_i})^{'}$ is a vector of n_i independent variables, θ_i and β_i are unknown intercept and slope, respectively, and $\boldsymbol{e}_i = (e_{i1}, \cdots, e_{in_i})^{'}$ is the vector of errors which are mutually independent and identically distributed as normal variable, that is, $\boldsymbol{e}_i \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_{n_i})$ where $\boldsymbol{I}_{n_i}$ is the identity matrix of order n_i. Equation (1.1) represents two linear models with different intercept and slope parameters. If $\beta_1 = \beta_2 = \beta$, then there are two parallel simple linear models when $\theta_i^{'}s$ are different.

This paper considers statistical tests with NSPI and the criteria that are used to compare the performance of the UT, RT and PTT are the size and power of the tests. A statistical test that has a minimum size is preferable because it will give a smaller probability of the Type I error. Furthermore, a test that has maximum power is preferred over any other tests because it guarantees the highest probability of rejecting any false null hypothesis. A test that minimizes the size and maximizes the power is preferred over any other tests. In reality, the size of a test is fixed, and then the choice of the best test is based on its maximum power.

This study considers testing the equality of the two intercepts when the equality of slopes is suspected. For which we focus on three different scenarios on the slope parameters, and define three different tests:

(i) for the UT, let ϕ^{UT} be the test function and T^{UT} be the test statistic for testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_a : \boldsymbol{\theta} > \boldsymbol{\theta}_0$ when $\boldsymbol{\beta} = (\beta_1, , \beta_2)'$ is unspecified,

(ii) for the RT, let ϕ^{RT} be the test function and T^{RT} is the test statistic for testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_a : \boldsymbol{\theta} > \boldsymbol{\theta}_0$ when $\boldsymbol{\beta} = \beta_0 \mathbf{1}_2$ (fixed vector),

(iii) for the PTT, let ϕ^{PTT} be the test function and T^{PTT} be the test statistic for testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_a : \boldsymbol{\theta} > \boldsymbol{\theta}_0$ following a pre-test (PT) on the slope parameters. For the PT, let ϕ^{PT} be the test function for testing $H_0^* : \boldsymbol{\beta} = \beta_0 \mathbf{1}_p$ (a suspected constant) against $H_a^* : \boldsymbol{\beta} > \beta_0 \mathbf{1}_2$ to remove the uncertainty. If the H_0^* is rejected in the PT, then the UT is used to test the intercept, otherwise the RT is used to test H_0. Thus, the PTT on H_0 depends on the PT on H_0^*, and is a choice between the UT and RT.

The unrestricted maximum likelihood estimator or least square estimator of intercept and slope vectors, $\boldsymbol{\theta} = (\theta_1, \theta_2)'$ and $\boldsymbol{\beta} = (\beta_1, \beta_2)'$, are given as

$$\widetilde{\boldsymbol{\theta}} = \overline{\boldsymbol{Y}} - \boldsymbol{T}\widetilde{\boldsymbol{\beta}} \text{ and } \widetilde{\boldsymbol{\beta}} = \frac{(\boldsymbol{x}_{\mathrm{i}}'\boldsymbol{y}_{\mathrm{i}}) - (\frac{1}{\mathrm{n_i}})[\mathbf{1}_{\mathrm{i}}'\boldsymbol{x}_{\mathrm{i}}\mathbf{1}_{\mathrm{i}}'\boldsymbol{y}_{\mathrm{i}}]}{n_{\mathrm{i}}\, Q_{\mathrm{i}}}, \tag{1.2}$$

where $\widetilde{\boldsymbol{\theta}} = (\widetilde{\theta}_1, \widetilde{\theta}_2)'$, $\widetilde{\boldsymbol{\beta}} = (\widetilde{\beta}_1, \widetilde{\beta}_2)'$, $\boldsymbol{T} = \mathrm{Diag}(\overline{x}_1, \overline{x}_2)$, $n_iQ_i = \boldsymbol{x}_i'\boldsymbol{x}_i - (\frac{1}{n_i})\left[\mathbf{1}_i'\boldsymbol{x}_i\right]$ and $\widetilde{\theta}_i = \overline{Y_i} - \widetilde{\beta}_i \overline{x}_i$ for $i = 1, 2$.

Furthermore, the likelihood ratio (LR) test statistic for testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_a : \boldsymbol{\theta} > \boldsymbol{\theta}_0$ is given by

$$F = \frac{\widetilde{\boldsymbol{\theta}}'\boldsymbol{H}'\boldsymbol{D}_{22}^{-1}\boldsymbol{H}\widetilde{\boldsymbol{\theta}}}{s_e^2}, \tag{1.3}$$

where $\boldsymbol{H} = \boldsymbol{I}_2 - \frac{1}{nQ}\mathbf{1}_2\mathbf{1}_2'\boldsymbol{D}_{22}^{-1}$, $\boldsymbol{D}_{22}^{-1} = \mathrm{Diag}(n_1 Q_1, \cdots, n_2 Q_2)$, $nQ = \sum_{i=1}^{2} n_i Q_i$, $n_iQ_i = \boldsymbol{x}_i'\boldsymbol{x}_i - \frac{1}{n_i}(\mathbf{1}_i'\boldsymbol{x}_i)^2$ and $s_e^2 = (n-4)^{-1}\sum_{i=1}^{p}(\boldsymbol{Y} - \widetilde{\boldsymbol{\theta}}_i\mathbf{1}_{n_i} - \widetilde{\boldsymbol{\beta}}\boldsymbol{x}_i)'(\boldsymbol{Y} - \widetilde{\boldsymbol{\theta}}_i\mathbf{1}_{n_i} - \widetilde{\boldsymbol{\beta}}\boldsymbol{x}_i)$ (Saleh, 2006, p. 14-15). Under H_0, F follows a central F distribution with $(1, n-4)$ degrees of freedom, and under H_a it follows a noncentral F distribution with $(1, n-4)$ degrees of freedom and noncentrality parameter $\boldsymbol{\Delta}^2/2$, where

$$\begin{aligned} \boldsymbol{\Delta}^2 &= \frac{\boldsymbol{\theta}'\boldsymbol{H}'\boldsymbol{D}_{22}^{-1}\boldsymbol{H}\boldsymbol{\theta}}{\sigma^2} = \frac{(\boldsymbol{\theta} - \boldsymbol{\theta}_0)'\boldsymbol{H}'\boldsymbol{D}_{22}^{-1}\boldsymbol{H}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)}{\sigma^2} \\ &= \frac{(\boldsymbol{\theta} - \boldsymbol{\theta}_0)'\boldsymbol{D}_{22}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)}{\sigma^2} \end{aligned} \tag{1.4}$$

and $\boldsymbol{D}_{22} = \boldsymbol{H}^{'}\boldsymbol{D}_{22}^{-1}\boldsymbol{H}$. When the slopes ($\boldsymbol{\beta}$) are equal to $\beta_0\mathbf{1}_2$ (specified), the restricted maximum likelihood estimator of the intercept and slope vectors are given as

$$\widehat{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\theta}} + \boldsymbol{T}\boldsymbol{H}\widetilde{\boldsymbol{\beta}} \text{ and } \widehat{\boldsymbol{\beta}} = \frac{\mathbf{1}_k\mathbf{1}_k^{'}\boldsymbol{D}_{22}^{-1}\widetilde{\boldsymbol{\beta}}}{n\boldsymbol{Q}}. \tag{1.5}$$

Section 2 provides the proposed three tests. Section 3 derives the distribution of the test statistics. The power function of the tests are obtained in Section 4. An illustrative example is given in Section 5. The comparison of the power of the tests and concluding remarks are provided in Sections 6 and 7.

2 The Proposed Tests

To test the equality of two intercepts when the equality of the slopes is suspected, we define three different test statistics as follows.

(i) For *unspecified* $\boldsymbol{\beta}$, the test statistic of the UT for testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_a : \boldsymbol{\theta} > \boldsymbol{\theta}_0$, under H_0, is given by

$$T^{UT} = \frac{\widetilde{\boldsymbol{\theta}}^{'}\boldsymbol{H}^{'}\boldsymbol{D}_{22}^{-1}\boldsymbol{H}\widetilde{\boldsymbol{\theta}}}{s_{ut}^2}, \tag{2.1}$$

where

$$s_{ut}^2 = (n-4)^{-1}\sum_{i=1}^{2}(\boldsymbol{Y} - \widetilde{\boldsymbol{\theta}}_i\mathbf{1}_{n_i} - \widetilde{\boldsymbol{\beta}}\boldsymbol{x}_i)^{'}(\boldsymbol{Y} - \widetilde{\boldsymbol{\theta}}_i\mathbf{1}_{n_i} - \widetilde{\boldsymbol{\beta}}\boldsymbol{x}_i).$$

The T^{UT} follows a central F distribution with $(1, n-4)$ degrees of freedom (d.f.). Under H_a, it follows a noncentral F distribution with $(1, n-4)$ d.f. and noncentrality parameter $\boldsymbol{\Delta}^2/2$. Under the normal model we have

$$\begin{pmatrix} \widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta} \\ \widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta} \end{pmatrix} \sim N_4\left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \ \sigma^2\begin{pmatrix} \boldsymbol{D}_{11} & -\boldsymbol{T}\boldsymbol{D}_{22} \\ -\boldsymbol{T}\boldsymbol{D}_{22} & \boldsymbol{D}_{22} \end{pmatrix}\right], \tag{2.2}$$

where $\boldsymbol{D}_{11} = \boldsymbol{N}^{-1} + \boldsymbol{T}\boldsymbol{D}_{22}\boldsymbol{T}\boldsymbol{\beta}$ and $\boldsymbol{N} = \mathrm{Diag}(n_1, \cdots, n_2)$.

(ii) For *specified* value of the slopes, $\boldsymbol{\beta} = \beta_0\mathbf{1}_2$ (fixed value), the test statistic of the RT for testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_a : \boldsymbol{\theta} > \boldsymbol{\theta}_0$ under H_0, is given by

$$T^{RT} = \frac{(\widehat{\boldsymbol{\theta}}^{'}\boldsymbol{H}^{'}\boldsymbol{D}_{22}^{-1}\boldsymbol{H}\widehat{\boldsymbol{\theta}}) + (\widetilde{\boldsymbol{\beta}}^{'}\boldsymbol{H}^{'}\boldsymbol{D}_{22}^{-1}\boldsymbol{H}\widetilde{\boldsymbol{\beta}})}{s_{rt}^2}, \tag{2.3}$$

where

$$s_{rt}^2 = (n-2)^{-1}\sum_{i=1}^{2}(\boldsymbol{Y} - \widehat{\boldsymbol{\theta}}_i \mathbf{1}_{n_i} - \widehat{\boldsymbol{\beta}}\boldsymbol{x}_i)'(\boldsymbol{Y} - \widehat{\boldsymbol{\theta}}_i \mathbf{1}_{n_i} - \widehat{\boldsymbol{\beta}}\boldsymbol{x}_i) \text{ and } \widehat{\boldsymbol{\beta}} = \beta_0 \mathbf{1}_2.$$

The T^{RT} follows a central F distribution with $(1, n-4)$ d.f.. Under H_a, it follows a noncentral F distribution with $(1, n-4)$ d.f. and noncentrality parameter $\boldsymbol{\Delta}^2/2$. Again, note that

$$\begin{pmatrix} \widehat{\boldsymbol{\theta}} - \boldsymbol{\theta} \\ \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} \end{pmatrix} \sim N_4 \left[\begin{pmatrix} \boldsymbol{T}\boldsymbol{H}\boldsymbol{\beta} \\ \mathbf{0} \end{pmatrix}, \ \sigma^2 \begin{pmatrix} \boldsymbol{D}_{11}^* & \boldsymbol{D}_{12}^* \\ \boldsymbol{D}_{12}^* & \boldsymbol{D}_{22} \end{pmatrix} \right], \tag{2.4}$$

where $\boldsymbol{D}_{11}^* = \boldsymbol{N}^{-1} + \frac{\boldsymbol{T}\mathbf{1}_2\mathbf{1}_2'\boldsymbol{T}\boldsymbol{\beta}}{nQ}$ and $\boldsymbol{D}_{12}^* = -\frac{1}{nQ}\mathbf{1}_2\mathbf{1}_2'\boldsymbol{T}$.

(iii) When the value of the slope is *suspected* to be $\boldsymbol{\beta} = \beta_0\mathbf{1}_2$ but unsure, a pre-test on the slope is required before testing the intercept. For the preliminary test (PT) of $H_0^* : \boldsymbol{\beta} = \beta_0 \mathbf{1}_p$ against $H_a^* : \boldsymbol{\beta} > \beta_0\mathbf{1}_2$, the test statistic under the null hypothesis is defined as

$$T^{PT} = \frac{\widetilde{\boldsymbol{\beta}}'\boldsymbol{H}'\boldsymbol{D}_{22}^{-1}\boldsymbol{H}\widetilde{\boldsymbol{\beta}}}{s_{ut}^2}, \tag{2.5}$$

which follows a central F distribution with $(1, n-4)$ d.f.. Under H_a, it follows a noncentral F distribution with $(1, n-4)$ d.f. and noncentrality parameter $\boldsymbol{\Delta}^2/2$. Again, note that

$$\begin{pmatrix} \widetilde{\boldsymbol{\theta}} - \beta_0\mathbf{1}_2 \\ \widetilde{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}} \end{pmatrix} \sim N_4 \left[\begin{pmatrix} (\widetilde{\beta^*} - \beta_0)\mathbf{1}_2 \\ \boldsymbol{H}\boldsymbol{\beta} \end{pmatrix}, \ \sigma^2 \begin{pmatrix} \mathbf{1}_2\mathbf{1}_2'/nQ & \mathbf{0} \\ \mathbf{0} & \boldsymbol{H}\boldsymbol{D}_{22} \end{pmatrix} \right], \tag{2.6}$$

where $\widetilde{\beta^*}\mathbf{1}_2 = \frac{\mathbf{1}_2\mathbf{1}_2'\boldsymbol{D}_{22}^{-1}\boldsymbol{\beta}}{nQ}$ (cf. Saleh, 2006, p. 273).

Let us choose a positive number α_j ($0 < \alpha_j < 1$, for j $= 1, 2, 3$) and real value F_{ν_1,ν_2,α_j} (with ν_1 as the numerator d.f. and ν_2 as the denominator d.f.) such that

$$P\left(T^{UT} > F_{1,n-4,\alpha_1} \mid \boldsymbol{\theta} = \boldsymbol{\theta}_0\right) = \alpha_1, \tag{2.7}$$

$$P\left(T^{RT} > F_{1,n-4,\alpha_2} \mid \boldsymbol{\theta} = \boldsymbol{\theta}_0\right) = \alpha_2, \tag{2.8}$$

$$P\left(T^{PT} > F_{1,n-4,\alpha_3} \mid \boldsymbol{\beta} = \beta_0\mathbf{1}_2\right) = \alpha_3. \tag{2.9}$$

Now the test function for testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_a : \boldsymbol{\theta} > \boldsymbol{\theta}_0$ is defined by

$$\Phi = \begin{cases} 1, & \text{if } \left(T^{PT} \le F_c,\, T^{RT} > F_b\right) \text{ or } \left(T^{PT} > F_c,\, T^{UT} > F_a\right); \\ 0, & \text{otherwise,} \end{cases} \tag{2.10}$$

where $F_a = F_{\alpha_1,1,n-4}$, $F_b = F_{\alpha_2,1,n-4}$ and $F_c = F_{\alpha_3,1,n-4}$.

3 Sampling Distribution of Test Statistics

To derive the power function of the UT, RT and PTT, the sampling distribution of the test statistics proposed in Section 2 are required. For the power function of the PTT the joint distribution of (T^{UT}, T^{PT}) and (T^{RT}, T^{PT}) is essential. Let $\{N_n\}$ be a sequence of local alternative hypotheses defined as

$$N_n : (\boldsymbol{\theta} - \boldsymbol{\theta}_0, \boldsymbol{\beta} - \beta_0 \mathbf{1}_2) = \left(\frac{\boldsymbol{\lambda}_1}{\sqrt{n}}, \frac{\boldsymbol{\lambda}_2}{\sqrt{n}} \right) = \boldsymbol{\lambda}, \tag{3.1}$$

where $\boldsymbol{\lambda}$ is a vector of fixed real numbers and $\boldsymbol{\theta}$ is the true value of the intercept. The local alternative is used only to compute the power of the tests for specific values of the parameters. Under N_n the value of $(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ is greater than zero and under H_0 the value of $(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ is equal zero.

Following Yunus and Khan (2011b) and equation (2.1), we define the test statistic of the UT when $\boldsymbol{\beta}$ is unspecified, under N_n, as

$$T_1^{UT} \quad = \quad T^{UT} - n \left\{ \frac{(\boldsymbol{\theta} - \boldsymbol{\theta}_0)' \boldsymbol{H}' \boldsymbol{D}_{22}^{-1} \boldsymbol{H} (\boldsymbol{\theta} - \boldsymbol{\theta}_0)}{s_{ut}^2} \right\}. \tag{3.2}$$

The T_1^{UT} follows a noncentral F distribution with noncentrality parameter which is a function of $(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ and $(1, n-4)$ d.f., under N_n.

From equation (2.3) under N_n, $(\boldsymbol{\theta} - \boldsymbol{\theta}_0) > 0$ and $(\boldsymbol{\beta} - \beta_0 \mathbf{1}_2) > 0$, the test statistic of the RT becomes

$$T_2^{RT} = T^{RT} - n \left\{ \frac{(\boldsymbol{\theta} - \boldsymbol{\theta}_0)' \boldsymbol{H}' \boldsymbol{D}_{22}^{-1} \boldsymbol{H} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) + (\boldsymbol{\beta} - \beta_0 \mathbf{1}_2)' \boldsymbol{H}' \boldsymbol{D}_{22}^{-1} \boldsymbol{H} (\boldsymbol{\beta} - \beta_0 \mathbf{1}_2)}{s_{rt}^2} \right\}. \tag{3.3}$$

The T_2^{RT} also follows a noncentral F distribution with a noncentrality parameter which is a function of $(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ and $(1, n-4)$ d.f., under N_n. Similarly, from the equation (2.5) the test statistic of the PT is given by

$$T_3^{PT} = T^{PT} - n \left\{ \frac{(\boldsymbol{\beta} - \beta_0 \mathbf{1}_2)' \boldsymbol{H}' \boldsymbol{D}_{22}^{-1} \boldsymbol{H} (\boldsymbol{\beta} - \beta_0 \mathbf{1}_2)}{s_{ut}^2} \right\}. \tag{3.4}$$

Under H_a, the T_3^{PT} follows a noncentral F distribution with a noncentrality parameter which is a function of $(\boldsymbol{\beta} - \beta_0 \mathbf{1}_2)$ and $(p-1, n-4)$ d.f..

From equations (2.1), (2.3) and (2.5) the T^{UT} and T^{PT} are correlated, and the T^{RT} and T^{PT} are uncorrelated. The joint distribution of the T^{UT} and T^{PT}, that is,

$$\begin{pmatrix} T^{UT} \\ T^{PT} \end{pmatrix}, \tag{3.5}$$

is a correlated bivariate F distribution with $(1, n-4)$ d.f.. The probability density function (pdf) and cumulative distribution function (cdf) of the correlated bivariate

F distribution is found in Krishnaiah (1964), Amos and Bulgren (1972) and El-Bassiouny and Jones (2009). Later, Johnson et al. (1995, p. 325) described a relationship of the bivariate F distribution with the bivariate beta distribution. This is due to the fact that the pdf of the bivariate F distribution has the same form as the pdf of the *beta distribution of the second kind.*

4 Power Function and Size of Tests

The power function of the UT, RT and PTT are derived below. From equation (2.1) and (3.2), (2.3) and (3.3), and (2.5), (2.10) and (3.4), the power function of the UT, RT and PTT are given, respectively, as

(i) the power of the UT

$$\begin{aligned}\pi^{UT}(\boldsymbol{\lambda}) &= P(T^{UT} > F_{\alpha_1,1,n-4} \mid N_n)\\ &= 1 - P\left(T_1^{UT} \le F_{\alpha_1,1,n-4} - \frac{\boldsymbol{\lambda}_1'\boldsymbol{H}'\boldsymbol{D}_{22}^{-1}\boldsymbol{H}\boldsymbol{\lambda}_1}{s_{ut}^2}\right)\\ &= 1 - P\left(T_1^{UT} \le F_{\alpha_1,1,n-4} - \frac{\boldsymbol{\lambda}_1'\boldsymbol{D}_{22}\boldsymbol{\lambda}_1}{s_{ut}^2}\right)\\ &= 1 - P\left(T_1^{UT} \le F_{\alpha_1,1,n-4} - k_{ut}\delta_1\right), \qquad (4.1)\end{aligned}$$

where $\delta_1 = \boldsymbol{\lambda}_1'\boldsymbol{D}_{22}\boldsymbol{\lambda}_1$ and $k_{ut} = \frac{1}{s_{ut}^2}$.

(ii) the power of the RT

$$\begin{aligned}\pi^{RT}(\boldsymbol{\lambda}) &= P\left(T^{RT} > F_{\alpha_1,1,n-4} \mid N_n\right)\\ &= P\left(T_2^{RT} > F_{\alpha_2,1,n-4} - \frac{(\boldsymbol{\theta}-\boldsymbol{\theta}_0)'\boldsymbol{H}'\boldsymbol{D}_{22}^{-1}\boldsymbol{H}(\boldsymbol{\theta}-\boldsymbol{\theta}_0)}{s_{rt}^2}\right)\\ &= 1 - P\left(T_2^{RT} \le F_{\alpha_2,1,n-4} - \frac{(\boldsymbol{\lambda}_1'\boldsymbol{H}'\boldsymbol{D}_{22}^{-1}\boldsymbol{H}\boldsymbol{\lambda}_1) + (\boldsymbol{\lambda}_2'\boldsymbol{H}'\boldsymbol{D}_{22}^{-1}\boldsymbol{H}\boldsymbol{\lambda}_2)}{s_{rt}^2}\right)\\ &= 1 - P\left(T_1^{RT} \le F_{\alpha_1,1,n-4} - k_{rt}(\delta_1+\delta_2)\right), \qquad (4.2)\end{aligned}$$

where $\delta_2 = \boldsymbol{\lambda}_2'\boldsymbol{D}_{22}\boldsymbol{\lambda}_2$ and $k_{rt} = \frac{1}{s_{rt}^2}$.

The power function of the PT is

$$\begin{aligned}\pi^{PT}(\boldsymbol{\lambda}) &= P\left(T^{PT} > F_{\alpha_3,1,n-4} | K_n\right)\\ &= 1 - P\left(T_3^{PT} \le F_{\alpha_3,1,n-4} - \frac{\boldsymbol{\lambda}_2'\boldsymbol{H}'\boldsymbol{D}_{22}^{-1}\boldsymbol{H}\boldsymbol{\lambda}_2}{s_{ut}^2}\right)\\ &= 1 - P\left(T_3^{PT} \le F_{\alpha_3,1,n-4} - k_{ut}\delta_2\right). \qquad (4.3)\end{aligned}$$

(iii) the power of the PTT

$$
\begin{aligned}
\pi^{PTT}(\boldsymbol{\lambda}) &= P\left(T^{PT} < F_{\alpha_3,1,n-4}, T^{RT} > F_{\alpha_2,1,n-4}\right) \\
&\quad + P\left(T^{PT} \geq F_{\alpha_3,1,n-4}, T^{UT} > F_{\alpha_1,1,n-4}\right) \\
&= \left(1 - \pi^{PT}\right)\pi^{RT} + d_{1r}\left(a, b\right), \qquad (4.4)
\end{aligned}
$$

where $d_{1r}(a,b)$ is bivariate F probability integral defined as

$$
\begin{aligned}
d_{1r}(a,b) &= \int_a^\infty \int_b^\infty f(F^{PT}, F^{UT}) dF^{PT} dF^{UT} \\
&= 1 - \int_0^a \int_0^b f(F^{PT}, F^{UT}) dF^{PT} dF^{UT}, \qquad (4.5)
\end{aligned}
$$

where

$$
a = F_{\alpha_3,1,n-4} - \frac{\boldsymbol{\lambda}_2' \boldsymbol{H}' \boldsymbol{D}_{22}^{-1} \boldsymbol{H} \boldsymbol{\lambda}_2}{(s_e^2} = F_{\alpha_3,1,n-4} - k_1\delta_2
$$

and

$$
b = F_{\alpha_1,1,n-4} - \frac{(\boldsymbol{\theta} - \boldsymbol{\theta}_0)' \boldsymbol{H}' \boldsymbol{D}_{22}^{-1} \boldsymbol{H} (\boldsymbol{\theta} - \boldsymbol{\theta}_0)}{s_e^2} = F_{\alpha_1,1,n-4} - k_1\delta_1.
$$

The integral $\int_0^a \int_0^b f(F^{PT}, F^{UT}) dF^{PT} dF^{UT}$ in equation (4.5) is the cdf of the correlated bivariate noncentral F distribution of the UT and PT. Following Yunus and Khan (2011c), we define the pdf and cdf of the bivariate noncentral F (BNCF) distribution, respectively, as

$$
\begin{aligned}
f(y_1, y_2) &= \left(\frac{m}{n}\right)^m \left[\frac{(1-\rho^2)^{\frac{m+n}{2}}}{\Gamma(n/2)}\right] \sum_{j=0}^\infty \sum_{r_1=0}^\infty \sum_{r_2=0}^\infty \left[\rho^{2j} \left(\frac{m}{n}\right)^{2j} \Gamma(m/2+j)\right] \\
&\quad \times \left[\left(\frac{e^{-\theta_1/2}(\theta_1/2)^{r_1}}{r_1!}\right) \left(\frac{\left(\frac{m}{n}\right)^{r_1}}{\Gamma(m/2+j+r_1)}\right) \left(y_1^{m/2+j+r_1-1}\right)\right] \\
&\quad \times \left[\left(\frac{e^{-\theta_2/2}(\theta_2/2)^{r_2}}{r_2!}\right) \left(\frac{\left(\frac{m}{n}\right)^{r_2}}{\Gamma(m/2+j+r_2)}\right) \left(y_2^{m/2+j+r_2-1}\right)\right] \\
&\quad \times \Gamma(q_{rj}) \left[(1-\rho^2) + \frac{m}{n} y_1 + \frac{m}{n} y_2\right]^{-(q_{rj})}, \text{ and} \qquad (4.6)
\end{aligned}
$$

$$
F(a,b) = P(Y_1 < a, Y_2 < b) = \int_0^a \int_0^b f(y_1, y_2) dy_1 dy_2, \qquad (4.7)
$$

where m is the numerator and n is the denominator degrees of freedom of the F variable. Setting $a = b = d$, Schuurmann et al. (1975) presented the critical values of d in a table of multivariate F distribution.

From equation (4.4), it is clear that the cdf of the BNCF distribution is involved in the expression of the power function of the PTT. Using equation (4.7), we evaluate the cdf of the BNCF distribution and use it in the calculation of the power function of the PTT. The relevant R codes are written, and the R package is used for the computation of the power and size and other graphical analyses.

Furthermore, the size of the UT, RT and PTT are given, respectively, as

(i) the size of the UT

$$\begin{aligned}\alpha^{UT} &= P\left(T^{UT} > F_{\alpha_1,1,n-4} \mid H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0\right) \\ &= 1 - P\left(T^{UT} \le F_{\alpha_1,1,n-4} \mid H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0\right) \\ &= 1 - P\left(T_1^{UT} \le F_{\alpha_1,1,n-4}\right), \end{aligned} \tag{4.8}$$

(ii) the size of the RT

$$\begin{aligned}\alpha^{RT} &= P\left(T^{RT} > F_{\alpha_2,1,n-4} \mid H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0\right) \\ &= 1 - P\left(T^{RT} \le F_{\alpha_2,1,n-4} \mid H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0\right) \\ &= 1 - P\left(T_2^{RT} \le F_{\alpha_2,1,n-4} - k_{rt}\delta_2\right). \end{aligned} \tag{4.9}$$

The size of the PT is given by

$$\begin{aligned}\alpha^{PT}(\boldsymbol{\lambda}) &= P\left(T^{PT} > F_{\alpha_3,1,n-4} | H_0\right) \\ &= 1 - P\left(T_3^{PT} \le F_{\alpha_3,1,n-4}\right). \end{aligned} \tag{4.10}$$

(iii) The size of the PTT

$$\begin{aligned}\alpha^{PTT} &= P\left(T^{PT} \le a, T^{RT} > d \mid H_0\right) + P\left(T^{PT} > a, T^{UT} > h \mid H_0\right) \\ &= P\left(T^{PT} < F_{\alpha_3,1,n-4}\right) P\left(T^{RT} > F_{\alpha_2,1,n-4}\right) + d_{1r}\left(a, h\right) \\ &= (1 - \alpha^{PT})\alpha^{RT} + d_{1r}\left(a, h\right), \end{aligned} \tag{4.11}$$

where $h = F_{\alpha_1,1,n-4}$.

5 A Simulation Example

For a simulation example we generated random data using R package (2013). Each of the two independent samples ($x_{ij},\ i = 1, 2,\ j = 1, \cdots, n_i$) were generated from the uniform distribution between 0 and 1. The errors ($e_i,\ i = 1, 2$) are generated from the normal distribution with $\mu = 0$ and $\sigma^2 = 1$. In each case $n_i = n = 100$ random variates were generated. The dependent variable (y_{1j}) was computed from the equation $y_{1j} = \theta_{01} + \beta_{11}x_{1j} + e_1$ for $\theta_{01} = 3$ and $\beta_{11} = 2$. Similarly, define $y_{2j} = \theta_{02} + \beta_{12}x_{2j} + e_2$ for $\theta_{02} = 3.6$ and $\beta_{12} = 2$, respectively. For the computation of the power function of the tests (UT, RT and PTT) we set $\alpha_1 = \alpha_2 = \alpha_3 = \alpha = 0.05$.

The graphs for the power function of the three tests are produced using the formulas in equations (4.1), (4.2) and (4.4). The graphs for the size of the three tests are produced using the formulas in equations (4.8), (4.9) and (4.11). The graphs of the power and size of the tests are presented in the Figures 1 and 2.

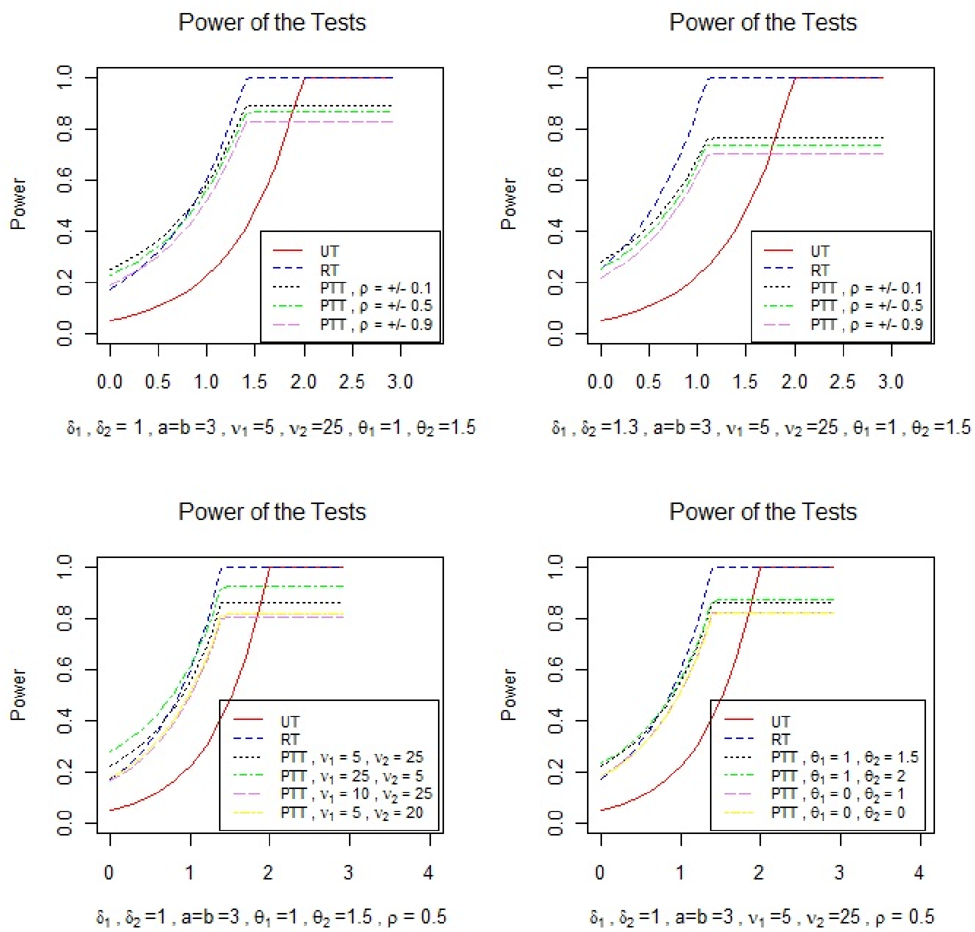

Figure 1: The power function of the UT, RT and PTT against δ_1 for some selected ρ, d.f. and noncentrality parameters.

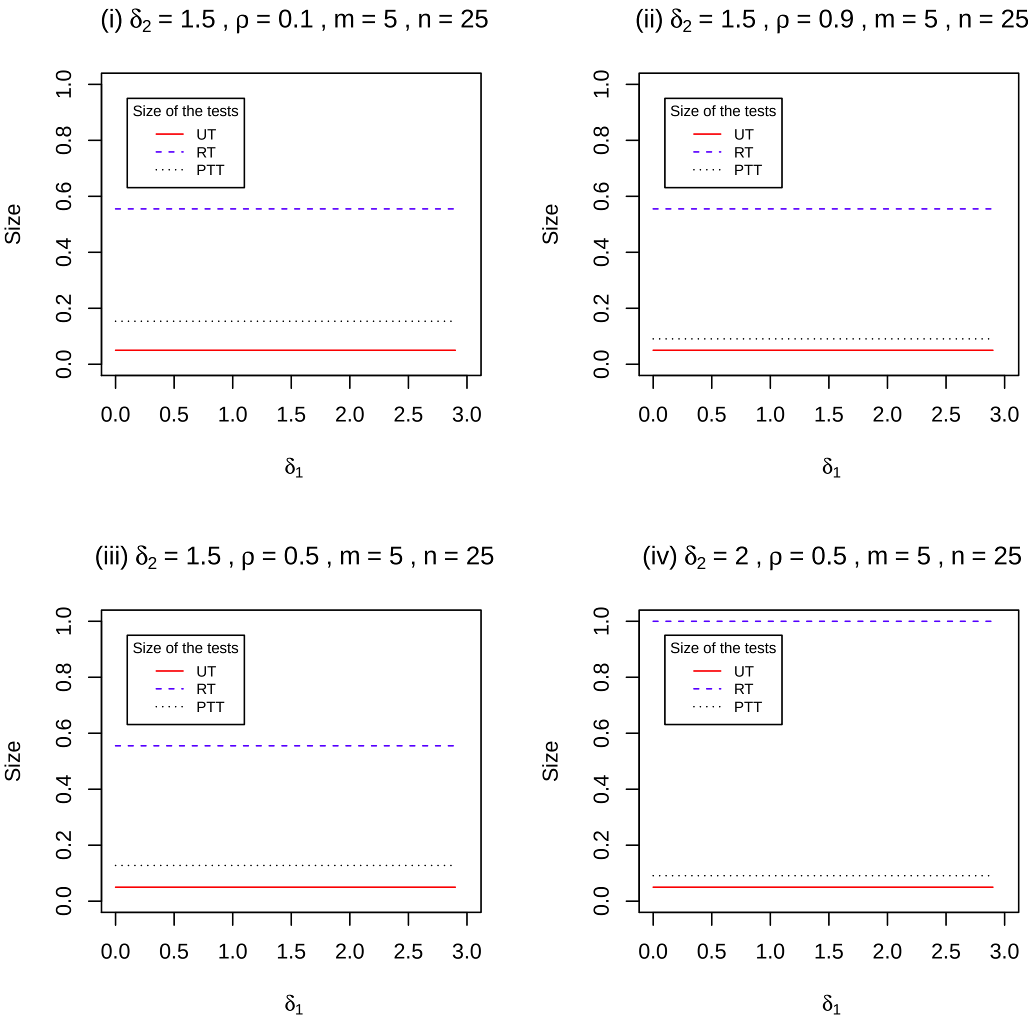

Figure 2: The size of the UT, RT and PTT against δ_1 for some selected ρ and δ_2.

6 Analyses of power and size

From Figure 1, as well as from equation (4.1), it is evident that the power of the UT does not depend on δ_2 and ρ, but it increases as the value of δ_1 increases. The form of the power curve of the UT is concave, starting from a very small value of near zero (when δ_1 is also near 0), it approaches 1 as δ_1 grows larger. The power of the UT increases rapidly as the value of δ_1 becomes larger. The minimum power of the UT is approximately 0.05 for $\delta_1 = 0$.

The shape of the power curve of the RT is also concave for all values of δ_1 and δ_2. The power of the RT increases as the values of δ_1 and/or δ_2 increase (see graphs in Figure 1(i) and 1(ii), and equation (4.2)). Moreover, the power of the RT is always larger than that of the UT for all values of δ_1 and/or δ_2. The minimum power of the RT is approximately 0.2 for $\delta_1 = 0$ and $\delta_2 = 0$. The maximum power of the RT is 1 for reasonably larger values of δ_1. The power of the RT reaches 1 much faster than that of the UT as δ_1 increases.

The power of the PTT depends on the values of δ_1, δ_2 and ρ (see Figure 1 and equation (4.4)). Like the power of the RT, the power of the PTT increases as the values of δ_1 increase. Moreover, the power of the PTT is always larger than that of the UT and RT for the values of δ_1 from around 0.7 to 1.5. The minimum power of the PTT is around 0.18 for $\delta_1 = 0$ (see Figure 1(i)), and it decreases as the value of δ_2 becomes larger. The gap between the power curves of the RT and PTT is much less than that between the UT and RT and/or UT and PTT. The power curve of the PTT tends to lie between the power curves of the UT and RT. However, the power of the PTT is identical for fixed value of ρ, regardless of its sign.

Figure 2 and equation (4.8) show that the size of the UT does not depend on δ_2. It is a constant and remains unchanged for all values of δ_1 and δ_2. The size of the RT increases as the value of δ_2 increases (see equation (4.7)). Moreover, the size of the RT is always larger than that of the UT. The size of the UT and RT do not depend on ρ.

The size of the PTT is closer to that of the UT for larger values of $\delta_2 > 2$. The difference (or gap) between the size of the RT and PTT increases significantly as the value of δ_2 and ρ increases. The size of the UT is $\alpha^{UT} = 0.05$ for all values of δ_1 and δ_2. For all values of δ_1 and δ_2, the size of the RT is larger than that of the UT, $\alpha^{RT} > \alpha^{UT}$. For all the values of ρ, $\alpha^{PTT} \leq \alpha^{RT}$. Thus, the size of the RT is always larger than that of the UT and PTT.

7 Concluding Remarks

Based on the analyses of the power for the three tests, the power of the RT is always higher than that of the UT for all values of δ_1 and δ_2. Also, the power of the PTT is always larger than that of the UT for all values δ_1 (see the curves on interval values of $0.7 < \delta_1 < 1.5$ for given simulated data), δ_2 and ρ.

For smaller values of δ_1 (see Figure 1) the PTT has higher power than the UT and RT. But for larger values of δ_1 the RT has higher power than the PTT and UT. Thus when the NSPI is reasonably accurate (that is δ_1 is small) the PTT over performs the UT and RT with higher power.

Since the size of the RT is the highest, and the PTT has larger size than UT, in terms of the size the UT is the best among the three tests. However, the UT performs the worst in terms of the power. Thus the PTT ensures higher power than the UT and lower size than the RT, and hence a better choice, especially when the NSPI on the slope parameters is reasonably accurate to be close to the true values.

The size of the PTT goes down as either the correlation coefficient (ρ) becomes larger (see graphs (i)-(ii) in Figure 2) or the value of δ_2 increases (see graphs (iii)-(iv) in Figure 2).

The extension of the work for testing one subset of the multiple regression model when NSPI is available on another subset is underway.

References

[1] Amos, D. E. and Bulgren, W. G. (1972). Computation of a multivariate F distribution. *Journal of Mathematics of Computation*, **26**, 255-264.

[2] Bancroft, T. A. (1944). On biases in estimation due to the use of the preliminary tests of singnificance. *Annals of Mathematical Statistics*, **15**, 190-204.

[3] Bancroft, T. A. (1964). Analysis and inference for incompletely specified models involving the use of the preliminary test(s) of singnificance. *Biometrics*, **20**(3), 427-442.

[4] El-Bassiouny, A. H. and Jones, M. C. (2009). A bivariate F distribution with marginals on arbitrary numerator and denominator degrees of freedom, and related bivariate beta and t distributions. *Statistical Methods and Applications*, **18**(4), 465-481.

[5] Han, C. P. and Bancroft, T. A. (1968). On pooling means when variance is unknown. *Journal of American Statistical Association*, **63**, 1333-1342.

[6] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). Continuous univariate distributions, Vol. 2, 2nd Edition. John Wiley and Sons, Inc., New York.

[7] Judge, G. G. and Bock, M. E. (1978). *The Statistical Implications of Pre-test and Stein-rule Estimators in Econoetrics.* North-Holland, New York.

[8] Khan, S. (2000). Improved estimation of the mean vector for Student-t model, Communications in Statistics-Theory and Methods, **29**(3), 507-527.

[9] Khan, S. (2003). Estimation of the Parameters of two Parallel Regression Lines Under Uncertain Prior Information. *Biometrical Journal*, **44**, 73-90.

[10] Khan, S. (2005). Estimation of parameters of the multivariate regression model with uncertain prior information and Student-t errors. *Journal of Statistical Research*, **39**(2), 79-94.

[11] Khan, S. (2006). Shrinkage estimation of the slope parameters of two parallel regression lines under uncertain prior information. *Journal of Model Assisted and Applications*,**1**, 195-207.

[12] Khan, S. (2008). Shrinkage estimators of intercept parameters of two simple regression models with suspected equal slopes. *Communications in Statistics - Theory and Methods*, **37**, 247-260.

[13] Khan, S. and Saleh, A. K. Md. E. (1997). Shrinkage pre-test estimator of the intercept parameter for a regression model with multivariate Student-t errors. *Biometrical Journal*, **39**, 1-17.

[14] Khan, S. and Saleh, A. K. Md. E. (2001). On the comparison of the pre-test and shrinkage estimators for the univariate normal mean. *Statistical Papers*, **42**(4), 451-473.

[15] Khan, S., Hoque, Z. and Saleh, A. K. Md. E. (2002). Estimation of the slope parameter for linear regression model with uncertain prior information, *Journal of Statistical Research*, **36**(1), 55-73.

[16] Khan, S. and Hoque, Z. (2003). Preliminary test estimators for the multivariate normal mean based on the modified W, LR and LM tests. *Journal of Statistical Research*, Vol 37, 43-55.

[17] Khan, S. and Saleh, A. K. Md. E. (2005). Estimation of intercept parameter for linear regression with uncertain non-sample prior information. *Statistical Papers.* **46**, 379-394.

[18] Khan, S. and Saleh, A. K. Md. E. (2008). Estimation of slope for linear regression model with uncertain prior information and Student-t error. *Communications in Statistics-Theory and Methods*, **37**(16), 2564-2581.

[19] Kleinbaum, D. G., Kupper, L. L., Nizam, A. and Muller, K. E. (2008). Applied regression analysis and other multivariable methods. Duxbury, USA.

[20] Krishnaiah, P. R. (1964). On the simultaneous anova and manova tests. Part of PhD thesis, University of Minnesota.

[21] R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

[22] Saleh, A. K. Md. E. (2006). Theory of preliminary test and Stein-type estimation with applications. John Wiley and Sons, Inc., New Jersey.

[23] Saleh, A. K. Md. E. and Sen, P. K. (1978). Nonparametric estimation of location parameter after a preliminary test on regression. *Annals of Statistics*, **6**, 154-168.

[24] Saleh, A. K. Md. E. and Sen, P. K. (1982). Shrinkage least squares estimation in a general multivariate linear model. *Procedings of the Fifth Pannonian Symposium on Mathematical Statistics*, 307-325.

[25] Schuurmann, F. J., Krishnaiah, P. R. and Chattopadhyay, A. K. (1975). Table for a multivariate F distribution. *The Indian Journal of Statistics* **37**, 308-331.

[26] Tamura, R. (1965). Nonparametric inferences with a preliminary test. *Bull. Math. Stat.* **11**, 38-61.

[27] Yunus, R. M. (2010). Increasing power of M-test through pre-testing. Unpublished PhD Thesis, University of Southern Queensland, Australia.

[28] Yunus, R. M. and Khan, S. (2007). Test for intercept after pre-testing on slope - a robust method. *In: 9th Islamic Countries Conference on Statistical Sciences (ICCS-IX): Statistics in the Contemporary World - Theories, Methods and Applications.*

[29] Yunus, R. M. and Khan, S. (2011a). Increasing power of the test through pre-test - a robust method. *Communications in Statistics-Theory and Methods*, **40**, 581-597.

[30] Yunus, R. M. and Khan, S. (2011b). M-tests for multivariate regression model. *Journal of Nonparamatric Statistics*, **23**, 201-218.

[31] Yunus, R. M. and Khan, S. (2011c). The bivariate noncentral chi-square distribution - A compound distribution approach. *Applied Mathematics and Computation*, **217**, 6237-6247.