Estimation of the Parameters of two Parallel Regression Lines Under Uncertain Prior Information

Shahjahan Khan
Department of Mathematics & Computing
University of Southern Queensland
Toowoomba, Queensland, Australia
Email: khans@usq.edu.au

Summary

The problem of parallelism for bi-linear regression lines arises in many real life investigations. For two linear regression models with normal errors, the estimation of the slope as well as the intercept parameters is considered when it is *apriori* suspected that the two lines are parallel. Three different estimators are defined by using both the sample data and the non-sample *uncertain prior information.* The relative performances of the *unrestricted, restricted and preliminary test* estimators are investigated based on the analysis of the bias, and risk functions under quadratic loss. An example based on a medical study is used to illustrate the method.

Key Words: Two parallel regression lines; non-sample uncertain prior information; multivariate normal distribution; central and non-central chi-squared and F-distributions; maximum likelihood, restricted and preliminary test estimators; bias and quadratic risk.

AMS 1991 Subject Classification: Primary 62F30 and Secondary 62J05.

1 Introduction

The linear regression method has a very wide range of real life applications. This popular and simple statistical method has been used in statistical analysis in almost every sphere of modern life. Customarily, the regression parameters are estimated by using the sample data alone. However, it is well known that the inclusion of non-sample prior information in the estimation of parameters is likely to improve the quality of the estimator in terms of good statistical properties. Bancroft (1944) first introduced the idea of preliminary test estimator. Such an estimator uses both the sample data and non-sample prior information in the form of a suspected null hypothesis. Appropriate statistical test is performed to remove the element of uncertainty in the null hypothesis. Then the preliminary test estimator is defined as a function of the sample data, the non-sample prior information and the test statistic. The idea can be applied to the parallelism problem with two regression equations, when it is apriori suspected that the slopes of the two regression lines are equal, but not sure. In this paper we define and investigate three different estimators of the intercept and the slope parameters of two linear regression lines by using the sample data as well as the non-sample uncertain prior information. The properties of the three different estimators are investigated through detailed analysis of the bias function and quadratic risk functions.

Consider a clinical study where the experimenter has collected two different data sets on the effect of two drugs for building two separate regression models. Alternatively, consider a sociologist or psychologist who has constructed two regression equations, one set for the males and another for the females. In both cases it may be useful to get some insight into whether or not the parameters of the two different regression models differ significantly across the two data sets. Moreover, the researcher may wish to combine the two data sets to formulate an overall regression model, if the respective parameters of the two different regression models do not differ significantly. However, in practical problems, the parameters of the models are usually unknown and the equality can only be suspected. This kind of suspicion may be treated as non-sample *uncertain prior information* and can be incorporated in the estimation of the parameters of the models.

To formulate the problem, consider the following two regression equations:

$$y_{1j} = \theta_1 + \beta_1 x_{1j} + e_{1j};\ \ j = 1, 2, \cdots, n_1 \text{ and } y_{2j} = \theta_2 + \beta_2 x_{2j} + e_{2j}; j = 1, 2, \cdots, n_2 \quad (1.1)$$

for the two data sets: $\boldsymbol{y} = [\boldsymbol{y}_1',\ \boldsymbol{y}_2']$ and $\boldsymbol{x} = [\boldsymbol{x}_1',\ \boldsymbol{x}_2']$ where $\boldsymbol{y}_1 = [y_{11}, y_{12}, \cdots, y_{1n_1}]'$, $\boldsymbol{y}_2 = [y_{21}, y_{22}, \cdots, y_{2n_2}]'$, $\boldsymbol{x}_1 = [x_{11}, x_{12}, \cdots, x_{1n_1}]'$ and $\boldsymbol{x}_2 = [x_{21}, x_{22}, \cdots, x_{2n_2}]'$. Note that y_{ij} is the j^{th} response of the i^{th} model and e_{ij} is the associated error component; x_{ij} is the j^{th} value of the regressor in the i^{th} model; and β_i and θ_i are the slope and intercept parameters of the i^{th} regression equation, for $i = 1, 2$. We assume that the errors are identically and independently distributed as normal variables. Our problem is to estimate the vector of intercept parameters, $\boldsymbol{\theta} = (\theta_1, \theta_2)'$, and that of the slope parameters, $\boldsymbol{\beta} = (\beta_1, \beta_2)'$, when equality of slopes is suspected, but not sure. The non-sample information of suspected equality of the slopes of the two regression equations as well as the sample data are used to estimate the parameters of the suspected parallelism model.

The two regression equations can be combined in a single model as

$$\boldsymbol{y} = X\boldsymbol{\Phi} + \boldsymbol{e} \quad (1.2)$$

where $\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{pmatrix}$, $X = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \boldsymbol{x}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \boldsymbol{x}_2 \end{pmatrix}$, $\boldsymbol{\Phi} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \beta_1 \\ \beta_2 \end{pmatrix}$ and $\boldsymbol{e} = \begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \end{pmatrix}$. Now, if it is suspected that the two lines are concurrent then the suspicion in the form of non-sample *uncertain prior information*, say β, then the null hypothesis becomes,

$$H_0 : \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{\Phi} = \begin{pmatrix} \beta \\ \beta \end{pmatrix}. \quad (1.3)$$

In general, the null hypothesis of equality of slopes is given by $H_0 : \ C\boldsymbol{\Phi} = \boldsymbol{r}$, and the alternative hypothesis, H_a : negation of the H_0, where C is a matrix and $\boldsymbol{\Phi}$ and $\boldsymbol{r}$ are vectors of appropriate orders. It is under the general null hypothesis in (1.3), we wish to estimate the slope and intercept parameters of the regression lines represented in (1.1).

The problem under consideration falls in the realm of statistical problems known as inference in the presence of *uncertain prior information.* The usual practice in the literature is to treat such *uncertain prior information* specified by H_0 as a "nuisance parameter". Then the uncertainty in the form of the "nuisance parameter" is removed by 'testing it out'. In a series of papers Bancroft (1944, 1964, 1972) addressed the problem, and proposed the well known *preliminary test* estimator. A host of other authors, notably Kitagawa (1963), Han and Bancroft (1968), Saleh and Han (1990), Ali and Saleh (1990), and Mahdi et al. (1998) contributed in the development of the method under the normal theory. Furthermore, Saleh and Sen (e.g., 1978, 1985) published a series of articles in this area exploring the nonparametric as well as the asymptotic theory based on the least square estimators. Bhoj and Ahsanullah (1993, 1994) discussed the problem of estimation of conditional mean for simple regression model. Khan and Saleh (1997) discussed the problem of shrinkage pre-test estimation for the multivariate Student-t regression model.

In this paper, we define the maximum likelihood estimator (mle) of the elements of $\boldsymbol{\Phi}$ in (1.2) assuming that the errors are independently and identically distributed as normal variables with mean 0 and unknown variance, σ^2. Such an estimator is known as the *unrestricted estimator* (UE) of $\boldsymbol{\Phi}$. Then we define the *restricted estimator* (RE) of $\boldsymbol{\Phi}$ under the constraint of the H_0. Finally, we define the *preliminary test estimator* (PTE) of $\boldsymbol{\Phi}$ by using an appropriate test statistic that can be employed to test the null hypothesis. The main objective of the paper is to study the properties of the three different estimators, namely the UE, RE and PTE, for both the intercept and the slope parameters of the two suspected parallel regression lines. Also, we investigate the relative performances of the estimators under different conditions. The analysis of the performances of the estimators are provided that can be used as a basis to select a 'best' estimator in a given situation. The comparisons of the estimators are based on the criteria of unbiasedness and risk under quadratic loss, both analytically and graphically.

The *preliminary test* estimators (PTE) are defined as a function of the test statistic appropriate for testing the null hypothesis as well as the UE and RE. From the definition, it yields the *unrestricted* estimator (UE) if the null hypothesis is rejected at a pre-selected level of significance; otherwise it becomes the *restricted* estimator (RE). Therefore, the preliminary test estimator indeed gives us a choice between the two estimators, UE and RE. A better compromise between the two extremes has been discussed by Khan and Saleh (1995) which is based on a confidence coefficient, c $(0 < c < 1)$ as a measure of trust of the null hypothesis.

In the next section, we define three different estimators of the previously defined vectors of the slope and intercept parameters. Some important results, that are necessary for the computations of bias and risk of the estimators are discussed in section 3. The expressions for bias of the estimators and their analyses are provided in section 4. The performance comparison of the estimators of the slope and intercept parameters based on the quadratic risk criterion is discussed in section 5. Section 6 provides an example based on a set of clinical data. Some concluding remarks are included in section 7.

2 Formulation of the estimators

Assume that the error term, e_{ij} in (1.1) is independent and identically distributed as a normal variable with $E(e_{ij}) = 0$ and $Var(e_{ij}) = \sigma^2$ for $i = 1, 2$ and all j. Then the *unrestricted estimator* (UE) of β_i and θ_i are obtained by applying the method of maximum likelihood (or equivalently the least squares method) as

$$\tilde{\beta}_i = \sum_{j=1}^{n_i} \frac{(x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i)}{n_i Q_i}, \quad \tilde{\theta}_i = \bar{y}_i - \tilde{\beta}_i \bar{x}_i \tag{2.1}$$

where $\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$, $\bar{y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij}$ and $n_i Q_i = \sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2$ for $i = 1, 2$. Thus the *unrestricted estimator* (UE) of the vectors of the slope and intercept, $\boldsymbol{\beta} = (\beta_1,\ \beta_2)'$ and $\boldsymbol{\theta} = (\theta_1,\ \theta_2)'$ becomes

$$\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_1,\ \tilde{\beta}_2)', \quad \tilde{\boldsymbol{\theta}} = (\tilde{\theta}_1,\ \tilde{\theta}_2)' = \bar{\boldsymbol{y}} - \boldsymbol{T}\tilde{\boldsymbol{\beta}} \tag{2.2}$$

where $\bar{\boldsymbol{y}} = (\bar{y}_1,\ \bar{y}_2)'$ and $\boldsymbol{T} = \text{Diag}\{\bar{x}_1,\ \bar{x}_2\}$, a 2×2 diagonal matrix. When the null hypothesis of equality of slopes holds, then the *restricted estimator* (RE) of the slope parameter becomes

$$\hat{\beta} = \frac{1}{nQ}\sum_{i=1}^{2} n_i Q_i \tilde{\beta}_i \quad \text{with} \quad nQ = \sum_{i=1}^{2} n_i Q_i. \tag{2.3}$$

Then the *restricted estimator* (RE) of the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are defined as

$$\hat{\boldsymbol{\beta}} = \hat{\beta}\boldsymbol{l}_2 = (\hat{\beta}, \hat{\beta})', \quad \hat{\boldsymbol{\theta}} = \bar{\boldsymbol{y}} - \boldsymbol{T}\hat{\boldsymbol{\beta}} = \tilde{\boldsymbol{\theta}} + \boldsymbol{T}J\tilde{\boldsymbol{\beta}} \tag{2.4}$$

where $J = I_2 - \frac{\boldsymbol{l}_2\boldsymbol{l}_2'}{nQ} D_2^{-1}$ in which $D_2^{-1} = \text{Diag}\{n_1Q_1, n_2Q_2\}$, $\boldsymbol{l}_2$ is a 2-tuples of ones and I_2 is the identity matrix of order 2.

To remove the *uncertainty* in the null hypothesis we require to test the H_0 by using an appropriate test statistic. For the current problem, we consider the likelihood ratio test given by the following statistic

$$L_n = \frac{(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})' D_2^{-1} (\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})}{s^2} = \frac{(\tilde{\boldsymbol{\beta}}' J') D_2^{-1} (J\tilde{\boldsymbol{\beta}})}{s^2} \tag{2.5}$$

where $s^2 = \frac{1}{m}\sum_{i=1}^{2}\sum_{j=1}^{n_i}[(y_{ij} - \bar{y}_i) - \tilde{\beta}_i(x_{ij} - \bar{x}_i)]^2$ with $m = (n - 4)$ and the numerator can be expressed as

$$n_1Q_1\left\{\tilde{\beta}_1\left(1 - \frac{n_1Q_1}{nQ}\right) - \tilde{\beta}_2\frac{n_2Q_2}{nQ}\right\}^2 + n_2Q_2\left\{\tilde{\beta}_2\left(1 - \frac{n_1Q_1}{nQ}\right) - \tilde{\beta}_1\frac{n_1Q_1}{nQ}\right\}^2. \tag{2.6}$$

Under the null hypothesis, the above test statistic follows a central F-distribution with 1 and m degrees of freedom (D.F.). Let F_α denote the $(1-\alpha)^{th}$ quantile of an $F_{1,m}$ variable such that $(1-\alpha)\times 100\%$ area under the curve of the distribution is to the left of F_α. Then, the *preliminary test estimator* (PTE) of the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are defined as

$$\hat{\boldsymbol{\beta}}^{pt} = \hat{\boldsymbol{\beta}}I(L_n < F_\alpha) + \tilde{\boldsymbol{\beta}}I(L_n \geq F_\alpha), \quad \hat{\boldsymbol{\theta}}^{pt} = \bar{\boldsymbol{y}} - \boldsymbol{T}\hat{\boldsymbol{\beta}}^{pt} = \tilde{\boldsymbol{\theta}} + \boldsymbol{T}J\tilde{\boldsymbol{\beta}}I(L_n < F_\alpha) \tag{2.7}$$

where $I(A)$ denotes an indicator function of the set A. The PTE, defined above, is a convex combination of the UE and the RE, and depends on the random coefficient, $\zeta = I(L_n < F_\alpha)$ whose value is $(1-\alpha)$ when the null hypothesis is true. Also note that the PTE is a simple compromise between the UE and RE. At a given level of significance, the PTE may simply be either the UE or the RE depending on the rejection and acceptance of the null hypothesis respectively. Therefore, for large values of L_n the PTE becomes the UE and for smaller values of L_n the PTE turns out to be the RE. Obviously, the PTE is a function of the test statistic as well as the level of significance, α. Hence, the PTE may change its value with a change in the choice of α. Therefore, a search for an optimal value of α may be desirable. In this paper, the optimality of the level of significance is in the sense of miniminsing the maximum risk of an estimator. Methods are available in the literature that provide optimal α, (see Akaike (1972), for instance). Another fact about the PTE is that it does not allow smooth transition between the two extremes, the UE and RE. Khan and Saleh (1995) provided a *shrinkage preliminary test estimator* to overcome such a problem.

Since we have defined three different estimators for the slope and the intercept parameter, a natural question arises as to which estimator should be used, and why? The answer to the question requires to investigate the performances of the estimators under different conditions. To study the properties of the above estimators of the slope and intercept vectors, some essential results are provided in the next section.

3 Some Preliminaries

In this section, we provide some useful results that are instrumental to the computation of expressions for bias and risk under quadratic loss function for the three different estimators. Fist, observe that the joint distribution of $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\theta}}$ is multivariate normal with

$$\mathrm{E}\begin{pmatrix}\tilde{\boldsymbol{\theta}}\\ \tilde{\boldsymbol{\beta}}\end{pmatrix} = \begin{pmatrix}\boldsymbol{\theta}\\ \boldsymbol{\beta}\end{pmatrix} \text{ and covariance matrix, } \mathrm{Cov}\begin{pmatrix}\tilde{\boldsymbol{\theta}}\\ \tilde{\boldsymbol{\beta}}\end{pmatrix} = \sigma^2\begin{pmatrix}D_1 & D_{12}\\ D_{21} & D_2\end{pmatrix} \tag{3.1}$$

where $D_{12} = D_{21}' = -D_2\boldsymbol{T}$ and $D_1 = \frac{\mathrm{Cov}(\tilde{\boldsymbol{\theta}})}{\sigma^2} = \boldsymbol{\psi} + \boldsymbol{T}D_2\boldsymbol{T}'$ with $\boldsymbol{\psi} = \mathrm{Diag}\left\{\frac{1}{n_1}, \frac{1}{n_2}\right\}$.

Note that the matrix D_2 has been specified in the definition of J in equation (2.4). Also note that $JD_2J' = D_2$, $D_2J'\boldsymbol{T} = -D_{12}\boldsymbol{T} + \frac{\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}'}{nQ}$ with $\bar{\boldsymbol{x}} = (\bar{x}_1, \bar{x}_2)'$. Then, note that the joint distribution of the statistics, $\tilde{\boldsymbol{\beta}}$ and $J\tilde{\boldsymbol{\beta}}$ is multivariate normal with the mean vector,

$$\mathrm{E}\begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ J\tilde{\boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta} \\ J\boldsymbol{\beta} \end{pmatrix} \text{ and covariance matrix, } \mathrm{Cov}\begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ J\tilde{\boldsymbol{\beta}} \end{pmatrix} = \sigma^2 \begin{pmatrix} D_2 & D_{12}^* \\ D_{21}^* & D_2^* \end{pmatrix} \tag{3.2}$$

where $D_2^* = D_2 - \frac{\boldsymbol{l}_2\boldsymbol{l}_2'}{nQ}$. Therefore, marginally each of $\tilde{\boldsymbol{\beta}}$ and $J\tilde{\boldsymbol{\beta}}$ has a multivariate normal distribution with respective mean vector and covariance matrix. But the conditional expectation of the statistic $(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})$, given $J\tilde{\boldsymbol{\beta}}$, becomes $\mathrm{E}[(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) \mid J\tilde{\boldsymbol{\beta}}] = J\tilde{\boldsymbol{\beta}} - J\boldsymbol{\beta}$. In the next section, we derive the expressions of bias for the three previously defined estimators of the slope and intercept vectors of parameters.

4 The bias of estimators

First, the expressions for the bias of UE of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are obtained as

$$B_1(\tilde{\boldsymbol{\beta}}) = E(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \mathbf{0}, \text{ and } B_1(\tilde{\boldsymbol{\theta}}) = E(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}) = \mathbf{0}. \tag{4.1}$$

Thus both $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\theta}}$ are unbiased estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ respectively. This is a well-known property of the mle for normal models. The bias of the RE of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ is found to be

$$B_2(\hat{\boldsymbol{\beta}}) = E(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = -J\boldsymbol{\beta} = -\boldsymbol{\delta}, \text{ and } B_2(\hat{\boldsymbol{\theta}}) = E(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = \boldsymbol{T\delta} \tag{4.2}$$

where $\boldsymbol{\delta} = J\boldsymbol{\beta} = \boldsymbol{\beta} - \beta\boldsymbol{l}_2$, deviation of $\boldsymbol{\beta}$ from its value under H_0. Clearly, the RE is biased. The amount of bias becomes unbounded as $\boldsymbol{\delta} \to \infty$, that is, if the true value of $\boldsymbol{\beta}$ is far away from it's hypothesized value, $\beta\boldsymbol{l}_2$. On the other hand the bias is zero when the null hypothesis is true. The same comment applies for the bias of $\hat{\boldsymbol{\theta}}$. Thus unlike the UE, the RE is biased.

Finally, the bias expressions for the PTE is obtained as

$$B_3(\hat{\boldsymbol{\beta}}^{pt}) = E(\hat{\boldsymbol{\beta}}^{pt} - \boldsymbol{\beta}) = -\boldsymbol{\delta}G_{3,m}(l_\alpha; \Delta), \quad B_3(\hat{\boldsymbol{\theta}}^{pt}) = E(\hat{\boldsymbol{\theta}}^{pt} - \boldsymbol{\theta}) = \boldsymbol{T\delta}G_{3,m}(l_\alpha; \Delta) \tag{4.3}$$

where $\Delta = \frac{\boldsymbol{\delta}' D_2^{-1}\boldsymbol{\delta}}{\sigma^2}$, $l_\alpha = \frac{1}{3}F_\alpha$ and $G_{3,m}(l_\alpha; \Delta) = \int_{z=0}^{l_\alpha} f(z)dz$ in which Z has a non-central $F-$distribution. For the computational purposes, $G_{3,m}(l_\alpha; \Delta)$ can be written as

$$G_{3,m}(l_\alpha; \Delta) = \sum_{r=0}^{\infty} \frac{e^{-\frac{\Delta}{2}}(\frac{\Delta}{2})^r}{r!} IB_{q_\alpha}^1\left(\frac{3}{2} + r, \frac{m}{2}\right) \tag{4.4}$$

where $IB_{q_\alpha}^1\left(\frac{3}{2} + r, \frac{m}{2}\right)$ is the incomplete beta function ratio and $q_\alpha = \frac{m}{m + F_{1,m}(\alpha)}$. In the derivation of the bias expression for the PTE we use the result of Appendix B1 of Judge and Bock (1978) as well as the results in the previous section.

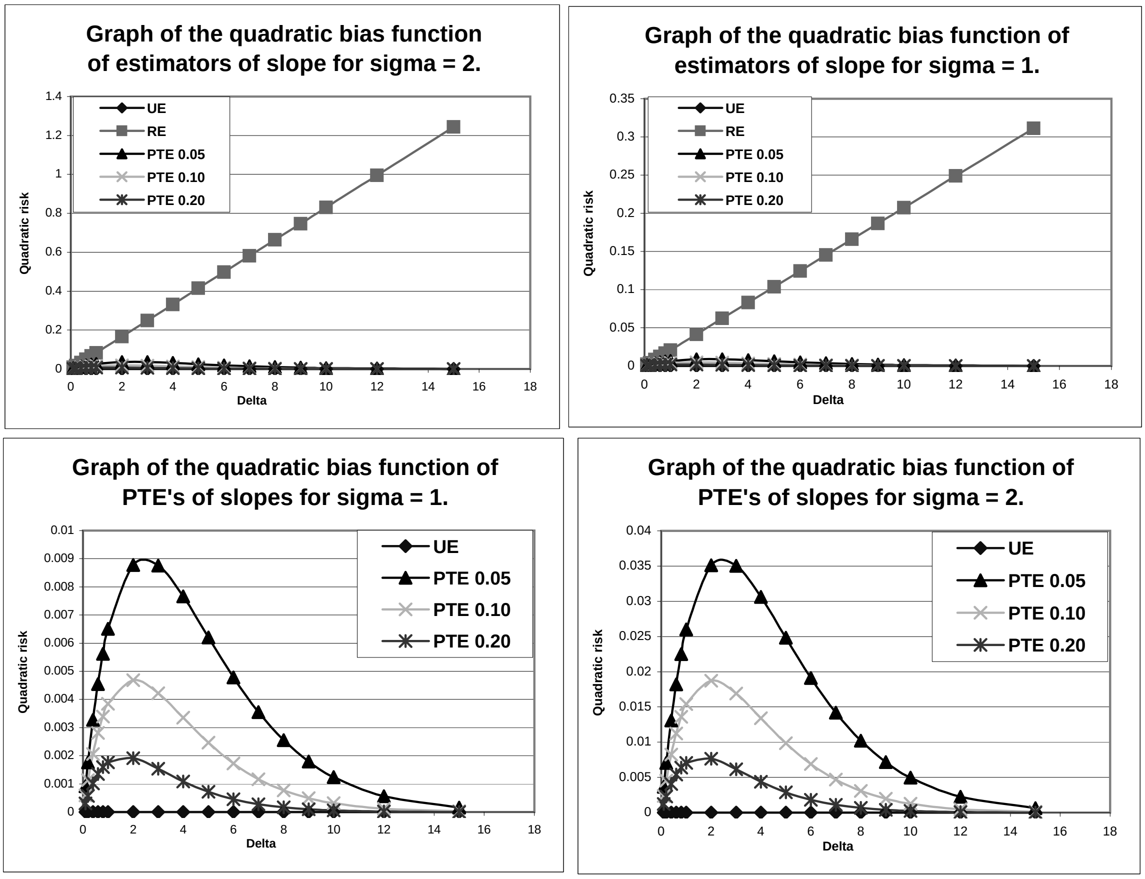

Figure 1: **Graph of Quadratic Bias functions of the estimators for** $\sigma = 1,\ 2$**.**

Obviously, the PTE is a biased estimator, and the amount of bias depends on the value of $G_{3,m}(\cdot)$, the non-central F distribution function and the extent of departure of the parameter from its value under null hypothesis. However, since $0 \leq G_{3,m}(\cdot) \leq 1$, the bias of the PTE is always smaller than that of the RE, except for $\Delta = 0$. This is true for both $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$.

The Quadratic Bias and Its Graph: The bias function of the slope vector as well as the intercept vector are also vectors of the same order. So direct comparison of the bias functions of the estimators are often not meaningful. To compare the overall bias of the estimators we define the quadratic bias as the vector product of the bias by itself. The quadratic bias is a scalar and it can be compared across the estimators. The plot of the quadratic bias function of the UE, RE and PTE with $\alpha = 0.05,\ 0.10$ and 0.20 are provided in Figure 1 for different values of the non-centrality parameter Δ. As expected, the quadratic bias of the UE is 0 for all values of Δ and that of the RE is unbounded and increases as the value of Δ grows large. The quadratic bias of the PTE is a function of the level of significance. As shown in the bottom two graphs in Figure 1, the shape of the curve of the quadratic bias function of the PTE is skewed to the right. At $\Delta = 0$ it starts for the origin and moves upward sharply until it reaches a pick for some moderate value of Δ and then gradually declines to the horizontal axis. The quadratic bias of the PTE increases as the preselected level of significance decreases. This is quite clear from the lower pair of graphs in Figure 1. The quadratic bias function of the RE and PTE increases as the variance of the population becomes larger.

5 The risk of estimators

For any estimator, $\boldsymbol{t}^*$ that estimates the parameter, $\boldsymbol{\mu}$, the quadratic error loss function is defined to be

$$L(\boldsymbol{t}^*, W, \boldsymbol{\mu}) = (\boldsymbol{t}^* - \boldsymbol{\mu})' W (\boldsymbol{t}^* - \boldsymbol{\mu})$$

where W is a positive definite matrix of appropriate dimension. Then the risk of $\boldsymbol{t}^*$ in estimating $\boldsymbol{\mu}$ is the expected value of $L(\boldsymbol{t}^*, W, \boldsymbol{\mu})$. Thus for the slope and intercept vectors, the quadratic risk functions are given by

$$R(\boldsymbol{\beta}^*, W_2, \boldsymbol{\beta}) = \mathrm{E}(\boldsymbol{\beta}^* - \boldsymbol{\beta})' W_2 (\boldsymbol{\beta}^* - \boldsymbol{\beta}) \text{ and } R(\boldsymbol{\theta}^*, W_1, \boldsymbol{\theta}) = \mathrm{E}(\boldsymbol{\theta}^* - \boldsymbol{\theta})' W_1 (\boldsymbol{\theta}^* - \boldsymbol{\theta}) \quad (5.1)$$

where $\boldsymbol{\beta}^*$ and $\boldsymbol{\theta}^*$ are the estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ respectively and W_1 and W_2 are two positive definite matrices of appropriate dimensions. Therefore, the expressions of the quadratic risk for the UE of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are obtained as

$$\begin{aligned} R_1(\tilde{\boldsymbol{\beta}}; W_2) &= E(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})' W_2 (\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \sigma^2 tr(W_2 D_2) \\ R_1(\tilde{\boldsymbol{\theta}}; W_1) &= E(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta})' W_1 (\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}) = \sigma^2 tr(W_1 D_1) \end{aligned} \quad (5.2)$$

respectively. Similarly, the risks of the RE of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are found to be

$$\begin{aligned} R_2(\hat{\boldsymbol{\beta}}; W_2) &= E(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})' W_2 (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \sigma^2 \frac{1}{nQ} tr(W_2 J^*) + \boldsymbol{\delta}' W_2 \boldsymbol{\delta} \\ R_2(\hat{\boldsymbol{\theta}}; W_1) &= E(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})' W_1 (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = \sigma^2 tr(W_1 D_{11}) + \boldsymbol{\delta}' \boldsymbol{T}' W_1 \boldsymbol{T} \boldsymbol{\delta} \end{aligned} \quad (5.3)$$

where $J^* = \boldsymbol{l}_2 \boldsymbol{l}_2'$. and $D_{11} = \Lambda + \frac{1}{nQ} \boldsymbol{t}\boldsymbol{t}'$ in which $\Lambda = \mathrm{Diag}\{\frac{1}{n_1}, \frac{1}{n_2}\}$ and $\boldsymbol{t}' = (\bar{x}_1, \bar{x}_2)$. Now, for the PTE, the quadratic risk expressions are given by

$$\begin{aligned} R_3(\hat{\boldsymbol{\beta}}^{pt}; W_2) &= E(\hat{\boldsymbol{\beta}}^{pt} - \boldsymbol{\beta})' W_2 (\hat{\boldsymbol{\beta}}^{pt} - \boldsymbol{\beta}) = \sigma^2 tr(W_2 D_2) \Big\{ 1 - G_{3,m}(l_\alpha; \Delta) \Big\} \\ &\quad + \boldsymbol{\delta}' W_2 \boldsymbol{\delta} \Big\{ 2 G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta) \Big\} \\ R_3(\hat{\boldsymbol{\theta}}^{pt}; W_1) &= E(\hat{\boldsymbol{\theta}}^{pt} - \boldsymbol{\theta})' W_1 (\hat{\boldsymbol{\theta}}^{pt} - \boldsymbol{\theta}) = \sigma^2 tr(W_1 D_{11}) \Big\{ 1 - G_{3,m}(l_\alpha; \Delta) \Big\} \\ &\quad + \boldsymbol{\delta}' \boldsymbol{T}' W_1 \boldsymbol{T} \boldsymbol{\delta} \Big\{ 2 G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta) \Big\}. \end{aligned} \quad (5.4)$$

The proof of the above results is straight forward by using the Appendix B1 of Judge and Bock (1978).

5.1 Risk analysis for estimators of slope

The comparisons of the risks are useful in studying the relative performances of the estimators and thereby selecting an appropriate estimator in a given situation. In this subsection we provide both the analytical and graphical analyses of the quadratic risk function of the estimators of the shape parameter.

Comparison of UE and RE

First consider the difference of the risks of the UE and the RE,

$$N_{12}(\tilde{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}; W_2) = R_1(\tilde{\boldsymbol{\beta}}; W_2) - R_2(\hat{\boldsymbol{\beta}}; W_2) = \sigma^2 tr(W_2 D_2) - \frac{\sigma^2}{nQ} tr(W_2 J^*) - \boldsymbol{\delta}' W_2 \boldsymbol{\delta}. \quad (5.5)$$

Thus the value of $N_{12}(\tilde{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}; W_2)$ is positive zero or negative depending on

$$\frac{\boldsymbol{\delta}' W_2 \boldsymbol{\delta}}{\sigma^2} \gtreqqless tr\Big(W_2\Big[D_2 - \frac{J^*}{nQ}\Big]\Big). \quad (5.6)$$

Therefore, the performance of the estimators depends on the value of $\boldsymbol{\delta}$. The RE over performs the UE if the actual value of the slope parameter is not far from its value under H_0. Otherwise, $\tilde{\boldsymbol{\beta}}$ dominates $\hat{\boldsymbol{\beta}}$. For further comparisons, note that by Courant Theorem (c.f. Puri and Sen, 1971, p.122) we have

$$\lambda_1 \leq \Big[\frac{\boldsymbol{\delta}' W_2 \boldsymbol{\delta}}{\boldsymbol{\delta}' D_2^{-1} \boldsymbol{\delta}}\Big] \leq \lambda_2 \quad (5.7)$$

where λ_1 is the smallest and λ_2 is the largest characteristic roots of the matrix $[W_2 D_2]$. Then we have $\Delta\lambda_1 \leq \Big[\frac{\boldsymbol{\delta}' W_2 \boldsymbol{\delta}}{\sigma^2}\Big] \leq \Delta\lambda_2$. Thus the risk of RE is bounded in the following way

$$R_1(\tilde{\boldsymbol{\beta}}; W_2) + \Delta\lambda_1 - \frac{\sigma^2}{nQ} tr(W_2 J^*) \leq R_2(\hat{\boldsymbol{\beta}}; W_2) \leq R_1(\tilde{\boldsymbol{\beta}}; W_2) + \Delta\lambda_2 - \frac{\sigma^2}{nQ} tr(W_2 J^*). \quad (5.8)$$

Clearly, when H_0 is true then $\Delta = 0$ and the bounds are equal. In a special case, if $W_2 = D_2^{-1}$ we get $\frac{tr(W_2 D_2)}{\lambda_2} = \frac{tr(W_2 D_2)}{\lambda_1} = 2$ and the difference between the risks becomes

$$N_{12}(\tilde{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}; W_2) \gtreqqless 0 \quad \text{according as} \quad \Delta \lesseqqgtr 2. \quad (5.9)$$

In another special case, if $W_2 = I_2$ then RE is superior to the UE if $\Delta \leq \frac{tr(W_2 D_2)}{\lambda_2}$, which depends on the value of the elements of the matrix D_2.

Comparison of UE and PTE

The risk-difference of the UE and the PTE is given by

$$\begin{aligned} N_{13}(\tilde{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}^{pt}; W_2) &= R_1(\tilde{\boldsymbol{\beta}}; W_2) - R_3(\hat{\boldsymbol{\beta}}^{pt}; W_2) = \sigma^2 tr(W_2 D_2) G_{3,m}(l_\alpha; \Delta) \\ &\quad - \boldsymbol{\delta}' W_2 \boldsymbol{\delta}\Big\{2G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta)\Big\}. \end{aligned} \quad (5.10)$$

Thus we have

$$N_{13}(\tilde{\boldsymbol\beta},\hat{\boldsymbol\beta}^{pt};W_2) \gtreqless 0 \quad \text{whenever} \quad \frac{\boldsymbol\delta' W_2 \boldsymbol\delta}{\sigma^2} \gtreqless \frac{tr(W_2D_2)G_{3,m}(l_\alpha;\Delta)}{\Big\{2G_{3,m}(l_\alpha;\Delta)-G_{5,m}(l_\alpha^*;\Delta)\Big\}}. \tag{5.11}$$

Then the bounds of $R_3(\hat{\boldsymbol\beta}^{pt};W_2)$ can be expressed as

$$R_3^L(\hat{\boldsymbol\beta}^{pt};W_2) \le R_3(\hat{\boldsymbol\beta}^{pt};W_2) \le R_3^U(\hat{\boldsymbol\beta}^{pt};W_2) \tag{5.12}$$

where

$$\begin{aligned} R_3^L(\hat{\boldsymbol\beta}^{pt};W_2) &= R_1(\hat{\boldsymbol\beta}^{pt};W_2)+\Delta\lambda_1\Big\{2G_{3,m}(l_\alpha;\Delta)-G_{5,m}(l_\alpha^*;\Delta)\Big\} \\ R_3^U(\hat{\boldsymbol\beta}^{pt};W_2) &= R_1(\hat{\boldsymbol\beta}^{pt};W_2)+\Delta\lambda_2\Big\{2G_{3,m}(l_\alpha;\Delta)-G_{5,m}(l_\alpha^*;\Delta)\Big\}. \end{aligned} \tag{5.13}$$

The bounds become equal when $\Delta = 0$, that is, when H_0 is true. But, under H_a

$$\begin{aligned} N_{13}(\tilde{\boldsymbol\beta},\hat{\boldsymbol\beta}^{pt};W_2) &\le 0 \quad \text{if} \quad \Delta \ge \frac{tr(W_2D_2)G_{3,m}(l_\alpha;\Delta)}{\lambda_1\Big\{2G_{3,m}(l_\alpha;\Delta)-G_{5,m}(l_\alpha^*;\Delta)\Big\}} \\ N_{13}(\tilde{\boldsymbol\beta},\hat{\boldsymbol\beta}^{pt};W_2) &\ge 0 \quad \text{if} \quad \Delta \le \frac{tr(W_2D_2)G_{3,m}(l_\alpha;\Delta)}{\lambda_2\Big\{2G_{3,m}(l_\alpha;\Delta)-G_{5,m}(l_\alpha^*;\Delta)\Big\}}. \end{aligned} \tag{5.14}$$

In a special case, if $W_2 = D_2^{-1}$ the difference between the risks becomes,

$$N_{13}(\tilde{\boldsymbol\beta},\hat{\boldsymbol\beta}^{pt};W_2) \gtreqless 0 \quad \text{according as } \Delta \lesseqgtr \frac{2G_{3,m}(l_\alpha;\Delta)}{\Big\{2G_{3,m}(l_\alpha;\Delta)-G_{5,m}(l_\alpha^*;\Delta)\Big\}}. \tag{5.15}$$

Furthermore, under H_0, $\Delta = 0$ and hence the risk of the PTE reduces to

$$R_3^0(\hat{\boldsymbol\beta}^{pt};W_2) = \sigma^2 tr(W_2D_2)\Big\{1-G_{3,m}(l_\alpha;0)\Big\} \tag{5.16}$$

which is less than that of the UE. But as Δ moves away from 0, the risk of the PTE increases and reaches a maximum at Δ_α (say) after crossing the line $\Delta_{0\alpha}$ given by (5.17) then decreases towards $\sigma^2 tr(W_2D_2)$, the risk of the UE as $\Delta \to \infty$.

Comparison of PTE and RE

The difference between the quadratic risks of the PTE and the RE is

$$\begin{aligned} N_{32}(\hat{\boldsymbol\beta}^{pt},\hat{\boldsymbol\beta};W_2) &= R_3(\hat{\boldsymbol\beta}^{pt};W_2)-R_2(\hat{\boldsymbol\beta};W_2) = \sigma^2 tr(W_2D_2)\Big\{1-G_{3,m}(l_\alpha;\Delta)\Big\} \\ &\quad -\boldsymbol\delta' W_2\boldsymbol\delta\Big\{1-2G_{3,m}(l_\alpha;\Delta)-G_{5,m}(l_\alpha^*;\Delta)\Big\}. \end{aligned} \tag{5.17}$$

Thus we get

$$N_{32}(\hat{\boldsymbol\beta}^{pt},\hat{\boldsymbol\beta};W_2) \gtreqless 0 \text{ according as } \frac{\boldsymbol\delta' W_2\boldsymbol\delta}{\sigma^2} \lesseqgtr \frac{tr(W_2D_2)\Big\{1-G_{3,m}(l_\alpha;\Delta)\Big\}}{\Big\{1-2G_{3,m}(l_\alpha;\Delta)-G_{5,m}(l_\alpha^*;\Delta)\Big\}}. \tag{5.18}$$

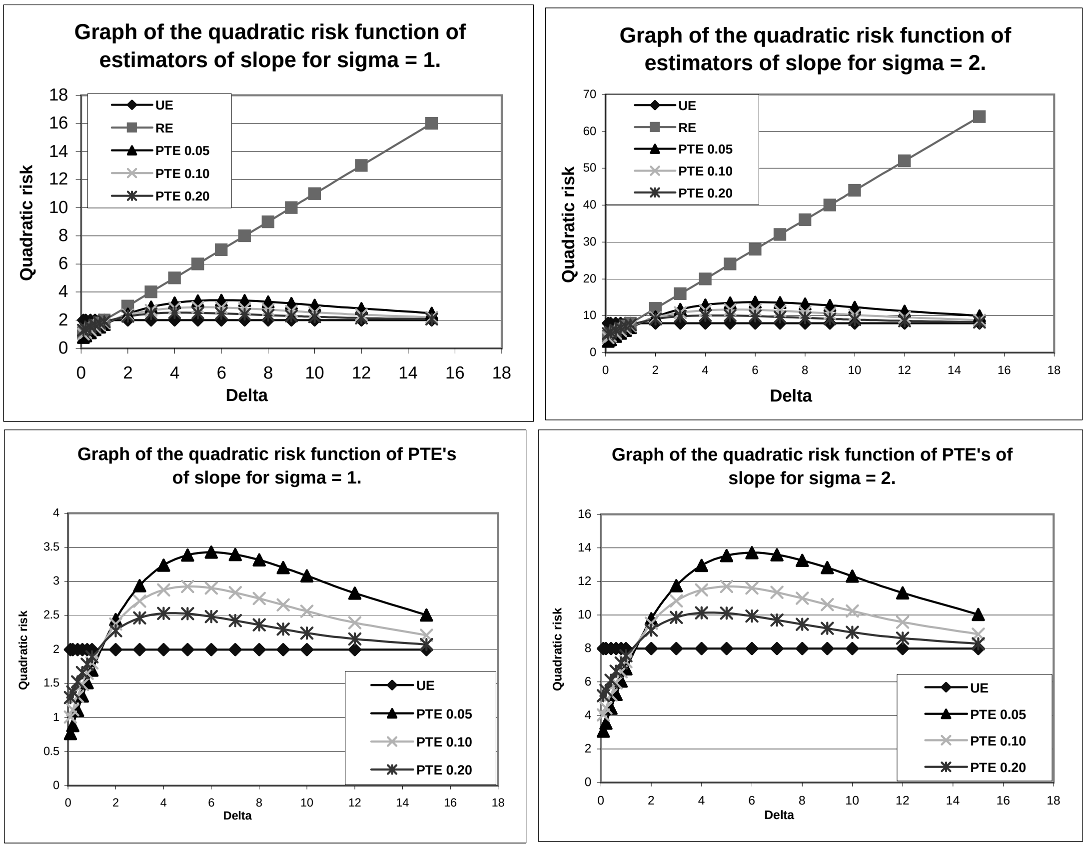

Figure 2: **Graph of Quadratic Risk functions of the estimators for $\sigma = 1,\ 2$.**

Therefore,

$$
\begin{aligned}
N_{32}(\hat{\boldsymbol{\beta}}^{pt}, \hat{\boldsymbol{\beta}}; W_2) &\geq 0 \text{ if } \Delta \leq \frac{tr(W_2D_2)\Big\{1 - G_{3,m}(l_\alpha; \Delta)\Big\}}{\lambda_1\Big\{1 - 2G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta)\Big\}} \text{ and} \\
N_{32}(\hat{\boldsymbol{\beta}}^{pt}, \hat{\boldsymbol{\beta}}; W_2) &\leq 0 \text{ if } \Delta \geq \frac{tr(W_2D_2)\Big\{1 - G_{3,m}(l_\alpha; \Delta)\Big\}}{\lambda_2\Big\{1 - 2G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta)\Big\}}.
\end{aligned} \tag{5.19}
$$

Under H_0, $\Delta = 0$ and hence the risk-difference reduces to $\sigma^2 tr(W_2D_2)\Big\{1 - G_{3,m}(l_\alpha; 0)\Big\}$, which is always positive. Thus the RE performs better than the PTE when H_0 is true. In a special case, when $W_2 = D_2^{-1}$,

$$
N_{32}(\hat{\boldsymbol{\beta}}^{pt}, \hat{\boldsymbol{\beta}}, W_2) \gtreqless 0 \quad \text{according as } \Delta \gtreqless \frac{2\Big\{1 - G_{3,m}(l_\alpha; \Delta)\Big\}}{\Big\{1 - 2G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta)\Big\}}. \tag{5.20}
$$

Graphical Analysis of Quadratic Risks: The graphs in Figure 2 display the risk functions of the estimators against the non-centrality parameter. The risk of the UE is constant and hence remains the same for all values of Δ. However, it increases with the increase in the value of σ. The quadratic risk of the RE is unbounded and increases as the value of Δ grows large. Nevertheless, it has smaller risk than the UE when the null hypothesis is true as well as when Δ is very small. But for larger values of Δ, the RE performs the worst. The quadratic risk function of the PTE depends on the selected level of significance. It is an inverse function of α for all Δ. When the null hypothesis is true, the PTE has a smallest risk among the three estimators, regardless of the value of α. This domination of the PTE continues up to some small value of Δ, (say Δ_P), and than the risk function of the PTE crosses that of the UE from the bottom and slowly grows up to maximum for some moderate value of Δ. Then it declines gradually towards the risk curve of the UE. The bottom two graphs show the behaviour of the PTE with the change of α and σ.

From the analytical results and graphical representation it is evident that there is no clear cut domination of one single estimator over the others for all values of Δ. If it is known that the null hypothesis is true, the RE is the best choice. But in real life, this is hardly the case. So, for unknown Δ, the RE could be the worst. The PTE is better than the UE is Δ is small or very large. For moderate values of Δ, the PTE is worse than the UE. This is more so when α is small.

5.2 Risk analysis for estimators of intercept

Finally, we compare the performances of the estimators of the intercept parameter vector based on the quadratic risk criterion.

Comparison of UE and RE

First consider the difference between the risks of the UE and the RE,

$$H_{12}(\tilde{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}; W_1) = R_1(\tilde{\boldsymbol{\theta}}; W_1) - R_2(\hat{\boldsymbol{\theta}}; W_1) = \sigma^2 tr(W_1 D_1) - \sigma^2 tr(W_1 D_{11}) - \boldsymbol{\delta}' \boldsymbol{T}' W_1 \boldsymbol{T} \boldsymbol{\delta}. \quad (5.21)$$

Thus the value of $H_{12}(\tilde{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}; W_1)$ is negative, zero or positive depending on

$$\Delta_T \gtreqless tr(W_1 D_{11}) - tr(W_1 D_1) = tr(W_1 \boldsymbol{T}' D_2 \boldsymbol{T}) - \frac{\bar{\boldsymbol{x}}' W_1 \bar{\boldsymbol{x}}}{nQ} \text{ with } \Delta_T = \frac{\boldsymbol{\delta}' \boldsymbol{T}' W_1 \boldsymbol{T} \boldsymbol{\delta}}{\sigma^2}. \quad (5.22)$$

Note that the matrix $[W_1 \boldsymbol{T}' D_2 \boldsymbol{T}] - \frac{\bar{\boldsymbol{x}}' W_1 \bar{\boldsymbol{x}}}{nQ}$ is positive semi-definite. Therefore, since $\boldsymbol{T}$ is not zero $tr(W_1[D_1 - D_{11}]) \geq 0$. From (5.27) it is evident that when $\boldsymbol{\delta}$ is close to zero RE performs better than the UE. On the other hand, as $\boldsymbol{\delta}$ moves away from zero $\boldsymbol{\delta}' \boldsymbol{T} W_1 \boldsymbol{T} \boldsymbol{\delta} \to \infty$, and hence the risk difference grows unboundedly. Then the UE performs better than the RE. Therefore, the UE is superior to the RE whenever

$$\frac{\boldsymbol{\delta}' \boldsymbol{T}' W_1 \boldsymbol{T} \boldsymbol{\delta}}{\sigma^2} > tr(W_1 \boldsymbol{T}' D_2 \boldsymbol{T}) - \frac{\bar{\boldsymbol{x}}' W_1 \bar{\boldsymbol{x}}}{nQ}. \quad (5.23)$$

Otherwise, the opposite conclusion holds.

Comparison of UE and PTE

The risk-difference of the UE and the PTE is given by

$$H_{13}(\tilde{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^{pt}; W_1) = R_1(\tilde{\boldsymbol{\theta}}; W_1) - R_3(\hat{\boldsymbol{\theta}}^{pt}; W_1) = \sigma^2 tr(W_1[D_1 - D_{11}])G_{3,m}(l_\alpha; \Delta) - \boldsymbol{\delta}'\boldsymbol{T}'W_1\boldsymbol{T}\boldsymbol{\delta}\Big\{2G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta)\Big\}. \quad (5.24)$$

Thus we have

$$H_{13}(\tilde{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}^{pt}; W_1) \gtreqless 0 \text{ whenever } \frac{\boldsymbol{\delta}'\boldsymbol{T}'W_1\boldsymbol{T}\boldsymbol{\delta}}{\sigma^2} \lesseqgtr \frac{tr(tr(W_1[D_1 - D_{11}])G_{3,m}(l_\alpha; \Delta)}{\Big\{2G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta)\Big\}}. \quad (5.25)$$

In a special case, when $W_1 = (D_1 - D_{11})^{-1}$ then (5.30) becomes

$$\frac{\boldsymbol{\delta}'\boldsymbol{T}'[D_1 - D_{11}]^{-1}\boldsymbol{T}\boldsymbol{\delta}}{\sigma^2} \lesseqgtr \frac{2G_{3,m}(l_\alpha; \Delta)}{\Big\{2G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta)\Big\}}. \quad (5.26)$$

Comparison of PTE and RE

The difference between the risks of the PTE and the RE is

$$H_{32}(\hat{\boldsymbol{\theta}}^{pt}, \hat{\boldsymbol{\theta}}; W_1) = R_3(\hat{\boldsymbol{\theta}}^{pt}; W_1) - R_2(\hat{\boldsymbol{\theta}}; W_1) = \sigma^2 tr(W_1[D_1 - D_{11}) \Big\{1 - G_{3,m}(l_\alpha; \Delta)\Big\} - \boldsymbol{\delta}'\boldsymbol{T}'W_1\boldsymbol{T}\boldsymbol{\delta}\Big\{1 - 2G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta)\Big\}. \quad (5.27)$$

Now, from (5.32) we get $H_{32}(\hat{\boldsymbol{\theta}}^{pt}, \hat{\boldsymbol{\theta}}; W_1) \lesseqgtr 0$ according as

$$\frac{\boldsymbol{\delta}'\boldsymbol{T}'W_1\boldsymbol{T}\boldsymbol{\delta}}{\sigma^2} \gtreqless \frac{tr(W_1[D_1 - D_{11}])\Big\{1 - G_{3,m}(l_\alpha; \Delta)\Big\}}{\Big[2\Big\{1 - 2G_{3,m}(l_\alpha; \Delta)\Big\} - \Big\{1 - G_{5,m}(l_\alpha^*; \Delta)\Big\}\Big]} \quad (5.28)$$

where $\Big\{1 - 2G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta)\Big\} = \Big[2\Big\{1 - 2G_{3,m}(l_\alpha; \Delta)\Big\} - \Big\{1 - G_{5,m}(l_\alpha^*; \Delta)\Big\}\Big]$. Therefore, based on (5.33), RE performs better than the PTE if

$$\Delta_T < \frac{tr(W_1[D_1 - D_{11}])\Big\{1 - G_{3,m}(l_\alpha; \Delta)\Big\}}{\lambda_2\Big[2\Big\{1 - 2G_{3,m}(l_\alpha; \Delta)\Big\} - \Big\{1 - G_{5,m}(l_\alpha^*; \Delta)\Big\}\Big]} \quad (5.29)$$

and the PTE dominates over the RE whenever

$$\Delta_T > \frac{tr(W_1[D_1 - D_{11}])\Big\{1 - G_{3,m}(l_\alpha; \Delta)\Big\}}{\lambda_1\Big[2\Big\{1 - 2G_{3,m}(l_\alpha; \Delta)\Big\} - \Big\{1 - G_{5,m}(l_\alpha^*; \Delta)\Big\}\Big]}. \quad (5.30)$$

The graphs in Figures 1 and 2, are produced for the quadratic bias and risk functions of the slope parameters. Similar graphs for the quadratic bias and risk functions can also be produced for the intercept parameters.

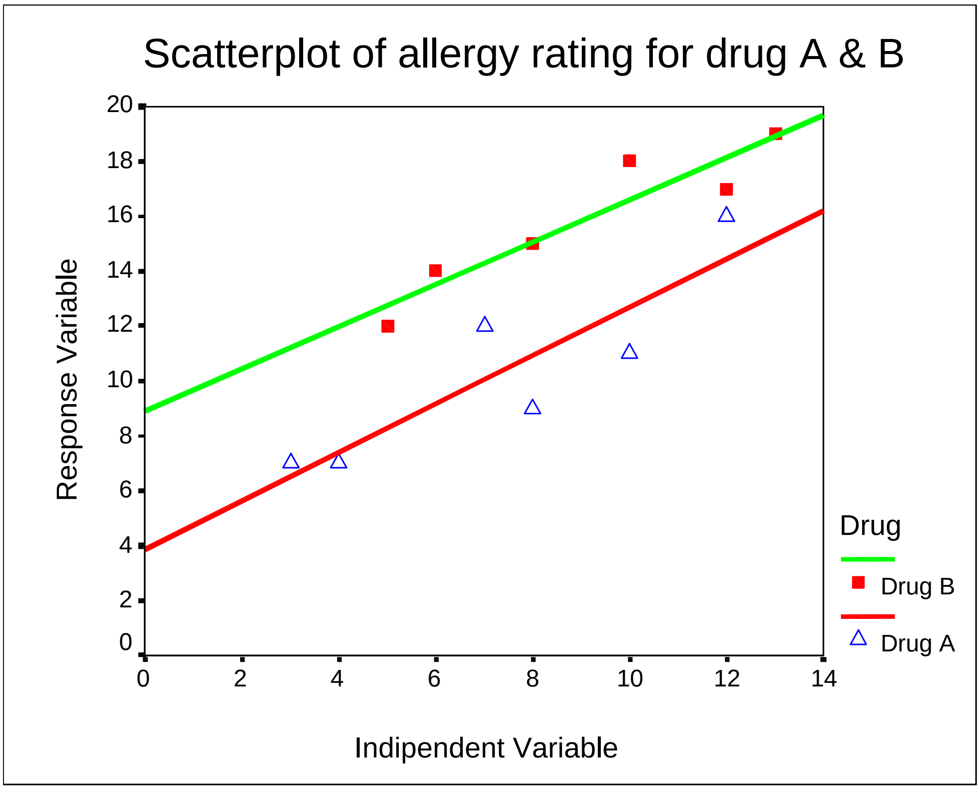

Figure 3: **Graph of two fitted regression lines for the allergy data.**

6 An example

To demonstrate the application of the method, we consider a data set from Weber and Skillings (2000, p.516). The study involves two drugs (A and B) for their effectiveness on testing allergies. Six people suffering from allergies were randomly allocated the the drugs, each one week apart. The severity of the allergy rated on a twenty-point scale before taking the drug, label as X, and after the drug, label as Y. Regression lines of Y on X have been fitted to the data for the two drugs separately. The scatterplot and the fitted regression lines are given in Figure 3. The fitted regression lines for the two data sets are

$$\hat{y}_A = 3.86 + 0.88x_A, \text{ and } \hat{y}_B = 8.91 + 0.77x_B. \tag{6.1}$$

Other statistics useful for the current study are $n_1Q_1 = 59.33$, $n_2Q_2 = 52.00$ and $nQ = 111.33$. The observed value of the test statistic is 0.3732 with a P-value of 0.3001. Hence there is not enough sample evidence to reject the null hypothesis of equal slopes, and thus the slopes of the two regression lines are not significantly different from one another.

6.1 Determination of optimal level of significance

The outcome of the preliminary test depends on the level of significance, so is the preliminary test estimator. Therefore, search for an optimal level of significance is obvious. Here the optimality of the level of significance is in the sense of miniminsing the maximum risk of an estimator. One method to obtain an optimal level of significance is to use the Akaike's (1972) Information Criterion or (AIC) as an abbreviation. Hirano (1977) used this approach to find the optimal level of significance. Khan and Saleh (1997) used the method in the linear regression model with Student-t errors.

For the model at hand we have 4 regression parameters and let the unrestricted parameter space be denoted by Ω. Under the null hypothesis, there are two regression parameters, and let the associated parameter space be denoted by Ω_0. Let the likelihood function under the unrestricted parameter space be denoted by $L(\Omega)$ and that under the null hypothesis be $L(\Omega_0)$. The corresponding AIC_Ω can be written as $-2log_e L(\tilde{\Omega}) + 2 \times 4$ and AIC_{Ω_0} can be written as $-2log_e L(\hat{\Omega}_0) + 2 \times 2$ respectively. Then following Hirano (1977), the AIC criterion for the model, $AIC_{\tilde{\Omega}} - AIC_{\hat{\Omega}_0} > 0$ turns our to be $-2log_e\lambda < 4$, where $\lambda = \frac{L(\tilde{\Omega})}{L(\hat{\Omega}_0)}$, in which, $L(\tilde{\Omega})$ is the unrestricted maximum of the likelihood function and $L(\hat{\Omega})$ is the maximum of the likelihood function under the null hypothesis. Since, for the current model, asymptotically $-2log_e\lambda$ follows a χ^2_2 distribution, the optimal level of significance based on the AIC criterion becomes $P(\chi^2_2 < 4) = 0.1353$. This optimal value of the level of significance can be used in the process of the preliminary test decision.

7 Concluding remarks

In this paper we have defined three different estimators for the slope and the intercept parameters of the two suspected parallel regression models. The performances of the three different estimators of the intercept and slope parameters have been analyzed by using the criteria of quadratic bias and risk under quadratic loss. The PTE has always smaller quadratic bias than the RE, except at $\Delta = 0$. But the quadratic bias of the UE is always 0 for all values of Δ. Based on the criterion of quadratic bias, the UE is the best among the three estimators. Based on the quadratic risk criterion, the superiority of estimators depends on various conditions discussed in section 5 and the graphs displayed in Figure 2. The RE is the best only if $\Delta = 0$. In the face of uncertainty on the value of Δ, if Δ is likely to be small then the PTE is the preferred option, regardless of the choice of α. One may use the UE as the best option if Δ is likely to be moderate, for which the quadratic risk of the PTE reaches its maximum. For very large values of Δ the PTE performs as good as the UE under the quadratic risk criterion. We have provided the marginal analysis of the problem. The joint study of the parameter sets of slopes and intercepts remains to be an open problem. Moreover, Stein-type shrinkage estimation is also possible for a set of $p > 2$

parallel regression models.

Acknowledgements

The author thankfully acknowledges some valuable suggestions of two unknown referees as well as Dr. M.I. Bhatti, that helped improve an earlier version of the paper.

References

Akaike, H. (1972). Information theory and an extention of maximum likelihood principle. Problems of Control and Informationn Theory. Academiai Kiado, Hungarian Academy of Science, 202-212.

Ali, A.M. and Saleh, A.K.Md.E. (1990). Estimation of the mean vector of a multivariate normal distribution under symmetry. *Jou. Statist. Comp. Simul.*, **35**, 209-226.

Bancroft, T.A. (1944). On biases in estimation due to the use of preliminary tests of significance. *Ann. Math. Statist.*, **15**, 190-204.

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

Bancroft, T.A. (1972). Some recent advances in inference procedures using preliminary tests of significance. In *Statistical Papers* in honour of G.W. Snedecor, Iowa State Univ. Press, 19-30.

Bhoj, D.S. and Ahsanullah, M. (1993). Estimation of conditional mean in a linear regression model. *Biometrical Journal*, **35**, 791-799.

Bhoj, D.S. and Ahsanullah, M. (1994). Estimation of a conditional mean in a regression model after a preliminary test on regression coeficient . *Biometrical Journal*, **36**, 153-163.

Han, C.P. and Bancroft, T.A. (1968). On pooling means when variance is unknown. *Jou. Amer. Statist. Assoc.*, **29**, 21-34.

Hirano, K. (1977). Estimation with procedures based on preliminary test, shrinkage technique and information criterion, *Ann. Inst. Statist. Math.*, Vol. 1, 361-379.

Judge, G.G. and M.E. Bock (1978). The statistical implication of pre-test and Stein-rule estimators in econometrics. North-Holland, New York.

Khan, S. and A.K.Md.E. Saleh (1997). Shrinkage pre-test estimator of the intercept parameter for a regression model with multivariate Student-t errors. *Biometrical Journal*, Vol. 29, 131-147.

Khan, S. and A.K.Md.E. Saleh (1995). Preliminary test estimators of the mean for sampling from some Student-t populations. *Journal of Statistical Research*, Vol. 29, 67-88.

Kitagawa, T. (1963). Estimation after preliminary tests of significance. University of California Publications in Statistics, **3**, 147-186.

Mahdi, T.N., Ahmad, S.E. and Ahsanullah, M. (1998). Improved predicitons: Pooling two identical regression lines. *Jou. Appld. Statist. Sc.*, **7**, 63-86.

Puri, M.L. and Sen, P.K. (1973). Nonparametric Methods in Multivariate Analysis. Wiley, New York.

Saleh, A.K.Md.E. and Han, C.P. (1990). Shrinkage estimation in regression analysis. *Estadistica,* **42**, 40-63.

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

Sen, P.K and Saleh, A.K.Md.E. (1985). On some shrinkage estimators of multivariate location. *Annals of Statistics.* **13**, 272-281.

Weber, D.C. and Skillings, J.H. (2000). A First Course of Design of Experiments: A Linear Model Approach. CRC Press, NY.