Shrinkage Estimation of the Slope 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

Abstract

The estimation of the slope parameter of two linear regression models with normal errors are considered, when it is *apriori* suspected that the two lines are parallel. The uncertain prior information about the equality of slopes is presented by a null hypothesis and a *coefficient of distrust* on the null hypothesis is introduced. The unrestricted estimator (UE) based on the sample responses and shrinkage restricted estimator (SRE) as well as shrinkage preliminary test estimator (SPTE) based on the sample responses and prior information are defined. The relative performances of the UE, SRE and SPTE are investigated based on the analysis of the bias, quadratic bias and quadratic risk functions. An example based on a health study data is used to illustrate the method. The SPTE dominates other two estimators if the *coefficient of distrust* is not far from 0 and the difference between the population slopes is small.

Keywords: Non-sample uncertain prior information; coefficient of distrust; maximum likelihood, shrinkage restricted and preliminary test estimators; analytical and graphical analysis; health study; bias and quadratic risk.

AMS 2000 Subject Classification: Primary 62F30 and Secondary 62J05.

1 Introduction

Traditionally the classical estimation methods exclusively use the sample data in the estimation of unknown parameters. Bancroft (1944) introduced the idea of inclusion of non-sample uncertain prior information in the estimation of the parameters. The method presents the non-sample uncertain prior information by a null hypothesis and removes the uncertainty through an appropriate statistical test, and is popularly known as the preliminary test estimator (PTE). Later, Stein (1956), and James and Stein (1961) introduced the well known James-Stein estimator as an improvement on the unrestricted estimator for multivariate models. They used the sample information as well as the non-sample information along with an appropriate test statistic in the definition of the estimator. The reason for the inclusion of non-sample information in conjunction with the sample information is to improve the statistical properties of the estimators. Recently, Khan and Saleh (2001) have used the coefficient of distrust $0 \leq d \leq 1$, a measure of degree of lack of trust on the null hypothesis, in the estimation of parameters. This coefficient of distrust reflects on the reliability of the prior information. In particular, $d = 0$ implies no distrust on the null hypothesis, $d = 0.5$ implies equal distrust and trust in the null hypothesis, and $d = 1$ implies total distrust in the null hypothesis. The selection of an appropriate value of d is subjective, and individual researcher would determine a specific value of d based on expert knowledge and, or, practical experiences. Combining the sample and non-sample information as well as the coefficient of distrust we propose the shrinkage restricted estimator (SRE) and shrinkage preliminary test estimator (SPTE), as a generalization of the restricted and preliminary test estimators, for the slope parameters of two suspected parallel linear regression models. Khan (2003) discussed different estimators of the slope under the suspected parallelism problem. However, it does not deal with the shrinkage preliminary test estimator.

The linear regression method is by far the most popular statistical tool that 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. The parallelism problem may arise in a variety of real world situation when there are two regression lines representing the same variables whose slopes are suspected to be equal. As an example, in the study of obesity among adult population, medical practitioners may be interested in the linear relationship between body fat and waist size by gender. In another example, the clinical researchers may wish to investigate the relationship between the systolic blood pressure rate and the age of smokers and non-smokers separately. In both cases the equality of the rate of change of the response variable on the explanatory variable could be suspected. Often, the researchers 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 of slopes can only be suspected with a certain degree of distrust. Each of the above cases can be modelled by the suspected parallelism of the pairs of regression lines. This kind of suspicion is treated as non-sample *uncertain prior information* and can be incorporated in the estimation of the parameters of the models.

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 specify such *uncertain prior information* by a H_0 and treat it 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), Mahdi et al. (1998), and Saleh (2006) 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 the next section, we introduce the parallelism model and define the null hypothesis to present the uncertain prior information. Section 3 defines three different estimators of the vector of the slope parameters. Some important results, that are necessary for the computations of bias and risk of the estimators are discussed in section 4. The expressions for bias of the estimators and their analyses are provided in section 5. The performance comparison of the estimators of the slope parameter based on the quadratic risk criterion is discussed in section 6. Section 7 provides an example based on a set of health study data. Some concluding remarks are included in section 8.

2 The Parallelism Problem

The parallelism problem can be described as a special case of two related regression lines on the same response variable. The explanatory variable is also the same, but coming from two different categories of the respondents. To formulate the problem, consider the following two regression equations:

$$y_{1j} = \theta_1 + \beta_1 x_{1j} + e_{1j};\ j = 1, 2, \cdots, n_1 \text{ and } y_{2j} = \theta_2 + \beta_2 x_{2j} + e_{2j}; j = 1, 2, \cdots, n_2 \quad (2.1)$$

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 $i = 1, 2$. Here $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ represent the same explanatory variable but coming from two different categories of respondents. In some cases a common set of responses may relate to two separate explanatory variables, but this study in not devoted to such cases. We assume that the errors are identically and independently distributed as normal variables with mean 0 and variance, σ^2 . Our problem is to estimate the vector of the slope parameters, $\boldsymbol{\beta} = (\beta_1, \beta_2)'$, when equality of slopes is suspected, but not sure. The non-sample prior 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. Furthermore, following Khan and Saleh (2001) the coefficient of distrust $0 \leq d \leq 1$ is introduced as a measure of the degree of lack of trust on the prior information.

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

$$\boldsymbol{y} = X\boldsymbol{\Phi} + \boldsymbol{e} \tag{2.2}$$

where $X = [X_1, X_2]'$ with $X_1 = [1, 0, x_1, 0]'$ and $X_2 = [0, 1, 0, x_2]'$, $\boldsymbol{\Phi} = [\theta_1, \theta_2, \boldsymbol{\beta}_1, \beta_2]'$ $\boldsymbol{e} = [\boldsymbol{e}_1, \boldsymbol{e}_2]'$, $\boldsymbol{\theta} = [\theta_1, \theta_2]'$ and $\boldsymbol{\beta} = [\beta_1, \beta_2]'$. Now, if it is suspected that the two lines are concurrent with common slope β then the suspicion in the non-sample *uncertain prior information*, can be expressed by the null hypothesis,

$$H_0 : \ [0, 0, 1, -1]\boldsymbol{\Phi} = 0. \tag{2.3}$$

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 known matrix and $\boldsymbol{r}$ is a known vector of appropriate order. It is under the suspected null hypothesis in (2.3), we wish to estimate the slope parameter of the regression lines represented in (2.1).

In this paper, we define the maximum likelihood estimator (mle) of the elements of $\boldsymbol{\beta}$ in (2.2) assuming that the errors are independent 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 $\boldsymbol{\beta}$. Then we define the shrinkage restricted estimator (SRE) of $\boldsymbol{\beta}$ under the constraint of the H_0 along with the associated degree of distrust. Finally, we define the shrinkage preliminary test estimator (SPTE) of $\boldsymbol{\beta}$ 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, SRE and SPTE, for the slope parameter 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.

Traditionally the PTEs are defined as a function of the test statistic appropriate for testing the null hypothesis as well as the UE and RE. In this paper we introduce the coefficient of distrust in the definition of the PTE. Thus, we define the SPTE as a linear combination of the UE and the SRE. Hence the SPTE depends on both the test statistic and the coefficient of distrust on the null hypothesis. 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 *shrinkage restricted* estimator (SRE). Therefore, the shrinkage preliminary test estimator indeed gives us a choice between the two estimators, UE and SRE.

3 Formulation of the estimators

From the specification of the model in (2.1), the *unrestricted estimator* (UE) of β_i is obtained by the method of maximum likelihood (or equivalently the least squares method) as

$$\tilde{\beta}_i = \sum_{j=1}^{n_i} \frac{(x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i)}{n_i Q_i} \tag{3.1}$$

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 $i = 1, 2$. Then the unrestricted estimator (UE) of the vector of the slopes $\boldsymbol{\beta} = (\beta_1,\ \beta_2)'$ becomes

$$\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_1,\ \tilde{\beta}_2)'. \tag{3.2}$$

When the null hypothesis of equality of slopes holds, then the *restricted estimator* (RE) of the slope parameter becomes

$$\hat{\beta} = \frac{1}{nQ}\sum_{i=1}^{2} n_i Q_i \tilde{\beta}_i \quad \text{with} \quad nQ = \sum_{i=1}^{2} n_i Q_i. \tag{3.3}$$

This is the pooled estimator of the slope parameter. Thus the *restricted estimator* (RE) of the slope vector $\boldsymbol{\beta}$ is defined as

$$\hat{\boldsymbol{\beta}} = \hat{\beta}\boldsymbol{l}_2 = (\hat{\beta}, \hat{\beta})' \tag{3.4}$$

where $\boldsymbol{l}_2 = [1, 1]'$. The *shrinkage restricted estimator* (SRE) of the slope vector is defined using the coefficient of distrust as well as the UE and RE of $\boldsymbol{\beta}$ as follows:

$$\hat{\boldsymbol{\beta}}_d = d\tilde{\boldsymbol{\beta}} + (1 - d)\hat{\boldsymbol{\beta}}. \tag{3.5}$$

Note that when $d = 0$ the SRE becomes the RE, and when $d = 1$ the SRE yields the UE. Thus the SRE is a convex combination of the UE and RE. Unlike the PTE, the SRE allows smooth transition between the UE and RE through different values of d.

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 based on the following statistic

$$L_n = \frac{(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})' D_3^{-1} (\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})}{s^2} \tag{3.6}$$

where $D_3^{-1} = \text{Diag}\{\frac{1}{n_1 Q_1} + \frac{1}{nQ}, \frac{1}{n_2 Q_2} + \frac{1}{nQ}\}$ and $s^2 = \frac{1}{m}\sum_{i=1}^{2}\sum_{j=1}^{n_i}[(y_{ij} - \bar{y}_i) - \tilde{\beta}_i(x_{ij} - \bar{x}_i)]^2$ with $m = (n - 4)$ and the numerator can be expressed as

$$(\tilde{\beta}_1 - \hat{\beta})^2 \frac{(n_1 Q_1 nQ)}{(n_1 Q_1 + nQ)} + (\tilde{\beta}_2 - \hat{\beta})^2 \frac{(n_2 Q_2 nQ)}{(n_2 Q_2 + nQ)}. \tag{3.7}$$

Under the null hypothesis, the above test statistic follows a central F-distribution with 2 and m degrees of freedom (D.F.). Let F_α denote the $(1 - \alpha)^{th}$ quantile of an $F_{2,m}$ 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 slope vector $\boldsymbol{\beta}$ is defined as

$$\hat{\boldsymbol{\beta}}^{pt} = \hat{\boldsymbol{\beta}} I(L_n < F_\alpha) + \tilde{\boldsymbol{\beta}} I(L_n \geq F_\alpha) \tag{3.8}$$

where $I(A)$ denotes an indicator function of the set A. The PTE, defined above, is a choice between the UE and the RE, and depends on the random coefficient, $\zeta = I(L_n < F_\alpha)$ whose value is 1 when the null hypothesis is accepted, and 0 otherwise. Also note that, unlike the SPTE (defined in eq 3.9), the PTE is a extreme choice 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 becomes essential. In this paper, the optimality of the level of significance is in the sense of minimising the maximum risk of an estimator. Methods are available in the literature that provide optimal α. Following Akaike (1972), Khan (2003) obtained an optimal value of α based on the AIC criterion . 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.

The *shrinkage preliminary test estimator* (SPTE) of the slope vector is defined as

$$\hat{\boldsymbol{\beta}}_d^{pt} = \hat{\boldsymbol{\beta}}_d I(L_n < F_\alpha) + \tilde{\boldsymbol{\beta}} I(L_n \geq F_\alpha). \tag{3.9}$$

A simpler form of the SPTE is expressed as

$$\hat{\boldsymbol{\beta}}_d^{pt} = \tilde{\boldsymbol{\beta}} - (1-d)(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}) I(L_n < F_\alpha). \tag{3.10}$$

Clearly the PTE is a special case of the SPTE. In particular, the SPTE becomes the PTE when $d = 0$, and it turns out to be the UE when $d = 1$, regardless of the value of $I(L_n \geq F_\alpha)$.

We have defined three different estimators for the slope parameter vector in this paper. 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 vector, some useful results are provided in the next section.

4 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.

First, observe that the joint distribution of the element of $\tilde{\boldsymbol{\beta}}$ is bivariate normal with

$$\mathrm{E}[\tilde{\boldsymbol{\beta}}] = \boldsymbol{\beta} \ \text{ and covariance matrix, } \mathrm{Cov}[\tilde{\boldsymbol{\beta}}] \ = \sigma^2 D_2 \tag{4.1}$$

where $D_2 = \mathrm{Diag}\{\frac{1}{n_1 Q_1}, \frac{1}{n_2 Q_2}\}$. Then, the joint distribution of the elements of $\hat{\boldsymbol{\beta}}$ is bivariate normal with the mean vector,

$$\mathrm{E}[\hat{\boldsymbol{\beta}}] \ = \boldsymbol{\beta}_0 = \beta \boldsymbol{l}_2 \text{ and covariance matrix, } \mathrm{Cov}[\hat{\boldsymbol{\beta}}] = \sigma^2 D_2^* \tag{4.2}$$

where $D_2^* = \mathrm{Diag}\{\frac{1}{nQ}, \frac{1}{nQ}\}$. Finally the distribution of $(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})$ is bivariate normal with

$$\mathrm{E}[\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}] = \boldsymbol{\delta} \ \text{ and covariance matrix, } \mathrm{Cov}[\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}] \ = \sigma^2 D_3 \tag{4.3}$$

where $\boldsymbol{\delta} = (\boldsymbol{\beta} - \boldsymbol{\beta}_0)$ and $D_3 = D_2 + D_2^* = \mathrm{Diag}\{\frac{1}{n_1 Q_1} + \frac{1}{nQ}, \frac{1}{n_2 Q_2} + \frac{1}{nQ}\}$.

In the next section, we derive the expressions of bias for the previously defined estimators of the slope parameters.

5 The bias of estimators

First, the expression for the bias of the UE of $\boldsymbol{\beta}$ is obtained as

$$B_1(\tilde{\boldsymbol{\beta}}) = E(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \mathbf{0}. \tag{5.1}$$

Thus $\tilde{\boldsymbol{\beta}}$ is an unbiased estimator of $\boldsymbol{\beta}$. This is a well-known property of the mle under the normal model. The bias of the RE of $\boldsymbol{\beta}$ is found to be

$$B_2^*(\hat{\boldsymbol{\beta}}) = E(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = -(\boldsymbol{\beta}_0 - \boldsymbol{\beta}) = -\boldsymbol{\delta} \tag{5.2}$$

where $\boldsymbol{\delta} = (\boldsymbol{\beta} - \beta \boldsymbol{l}_2)$, deviation of $\boldsymbol{\beta}$ from its value under H_0. Clearly, the RE is biased. The amount of bias of the RE 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. Thus unlike the UE, the RE is biased except then the null hypothesis is true.

The bias of the SRE of $\boldsymbol{\beta}$ is found to be

$$B_2(\hat{\boldsymbol{\beta}}_d) = E[\hat{\boldsymbol{\beta}}_d - \boldsymbol{\beta}] = -(1-d)\boldsymbol{\delta}. \tag{5.3}$$

The bias of the SRE becomes that of the RE when $d = 0$ and that of the UE when $d = 1$. Also, the bias expression for the PTE of the slope vector is obtained as

$$B_3^*(\hat{\boldsymbol{\beta}}^{pt}) = E(\hat{\boldsymbol{\beta}}^{pt} - \boldsymbol{\beta}) = -\boldsymbol{\delta} G_{3,m}(l_\alpha; \Delta) \tag{5.4}$$

where $\Delta = \frac{\boldsymbol{\delta}' D_3^{-1} \boldsymbol{\delta}}{\sigma^2}$, $l_\alpha = \frac{1}{3} F_\alpha$ and $G_{3,m}(l_\alpha; \Delta) = \int_{z=0}^{l_\alpha} f_Z(z) dz$ in which Z has a non-central $F-$distribution. For the computational purposes, $G_{3,m}(l_\alpha; \Delta)$ can be written as

$$G_{3,m}(l_\alpha; \Delta) = \sum_{r=0}^{\infty} \frac{e^{-\frac{\Delta}{2}} (\frac{\Delta}{2})^r}{r!} IB_{q_\alpha}^1 \left(\frac{3}{2} + r, \frac{m}{2} \right) \tag{5.5}$$

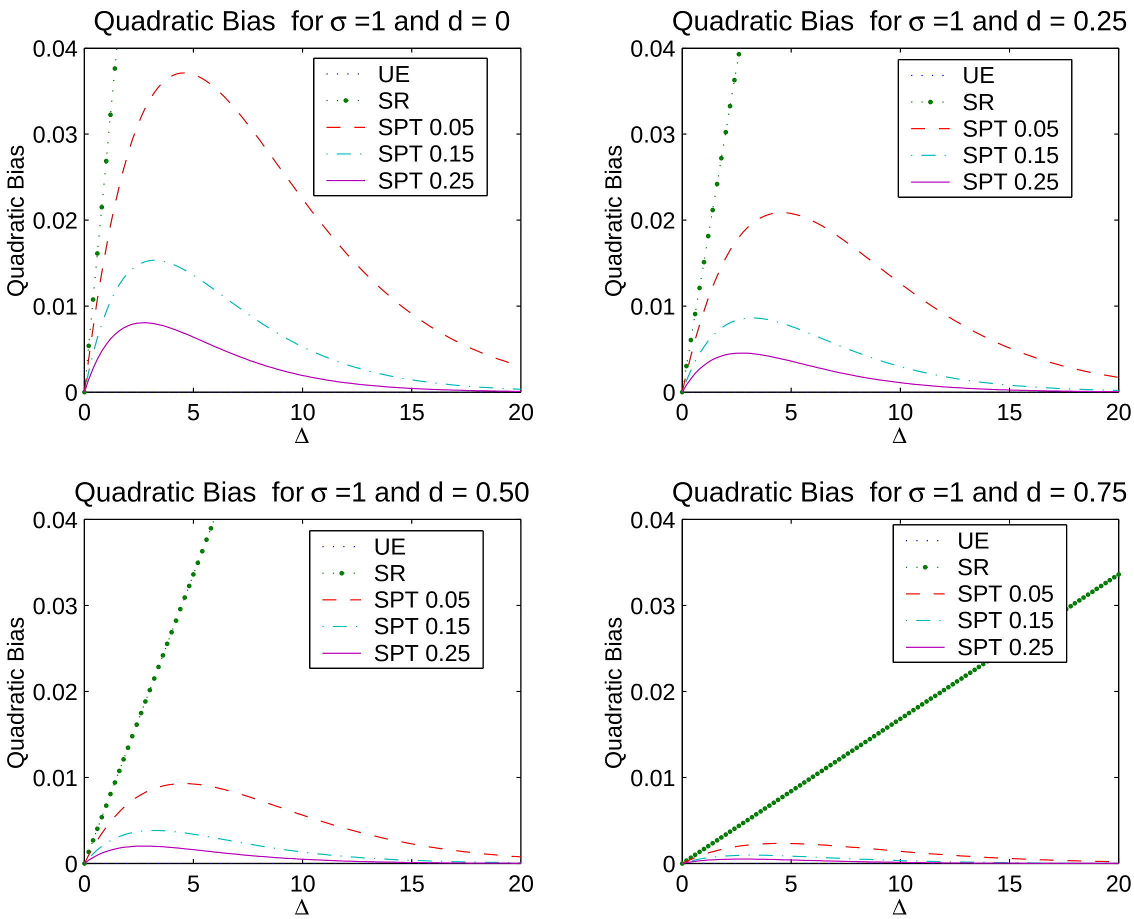

Figure 1: **Graph of Quadratic Bias functions of the estimators for Different Values of d When $\sigma^2 = 1$. Body Fat Example.**

where $IB^1_{q_\alpha}\left(\frac{3}{2}+r, \frac{m}{2}\right)$ is the incomplete beta function ratio with $q_\alpha = \frac{m}{m+F_{1,m}(\alpha)}$. 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.

Similarly, the bias of the SPTE of $\boldsymbol{\beta}$ is

$$B_3(\hat{\boldsymbol{\beta}}^{pt}_d) = E[\hat{\boldsymbol{\beta}}^{pt}_d - \boldsymbol{\beta}] = -(1-d)\boldsymbol{\delta}G_{3,m}(l_\alpha; \Delta). \tag{5.6}$$

The bias of the SPTE becomes that of the PTE when $d = 0$.

Obviously, the SPTE is a biased estimator, and the amount of bias depends on the value of $G_{3,m}(\cdot)$, the cdf of the non-central F distribution and the extent of departure of the parameter from its value under the null hypothesis. However, since $0 \le G_{3,m}(\cdot) \le 1$, the bias of the SPTE is always smaller than that of the SRE, except for $\Delta = 0$.

5.1 The Quadratic Bias and Its Graph

Obviously the bias function of the slope vector is also a vector of the same order. Therefore direct comparison of the bias functions of the estimators are not meaningful. So for the sake of comparing 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 quantity and hence it can be compared for various estimators. The plot of the quadratic bias function of the UE, SRE and SPTE with $\alpha = 0.05,\ 0.15$ and 0.25 are provided in Figure 1 for different values of the non-centrality parameter Δ and d when $\sigma = 1$. The quadratic bias of the UE is 0 for all values of Δ and that of the SRE is unbounded and increases as the value of Δ grows large regardless of the value of d. However, as the value of d approaches from 0 to 1, the growth rate of the quadratic bias of the SRE declines remarkably. At $d = 1$, the quadratic bias of the SRE is 0, as the SRE becomes the UE for that particular value of d.

The quadratic bias of the SPTE is a function of the level of significance as well as the coefficient of distrust. As shown in the graphs in Figure 1, the shape of the curve of the quadratic bias function of the SPTE is skewed to the right. At $\Delta = 0$ it starts at 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. This is true for all values of α and d, except at $d = 1$. Thus for very large values of Δ the quadratic bias of the SPTE is no different from that of the UE. The quadratic bias of the SPTE increases as the preselected level of significance decreases. This is quite clear from the graphs in Figure 1. At $d = 0$, the quadratic bias of the SPTE equals that of the PTE. But at $d = 1$, the quadratic bias of the SPTE becomes 0. Therefore when there is a complete distrust on the prior information then the quadratic bias of both the SRE and SPTE is no different from that of the UE.

The quadratic bias function of the SRE and SPTE increases as the variance of the population grows larger.

6 The risk of estimators

Let $\boldsymbol{t}^*$ be an estimator of the parameter, $\boldsymbol{\mu}$. Then the quadratic error loss function of $\boldsymbol{t}^*$ is defined as

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

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

$$R(\boldsymbol{\beta}^*, W_2, \boldsymbol{\beta}) = \mathrm{E}(\boldsymbol{\beta}^* - \boldsymbol{\beta})' W_2 (\boldsymbol{\beta}^* - \boldsymbol{\beta}) \tag{6.1}$$

where $\boldsymbol{\beta}^*$ is the estimator of $\boldsymbol{\beta}$ and W_2 is a positive definite matrix of appropriate dimensions. Therefore, the expression of the quadratic risk for the UE of $\boldsymbol{\beta}$ becomes

$$R_1(\tilde{\boldsymbol{\beta}}; W_2) \quad = \quad E(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})' W_2 (\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \sigma^2 tr(W_2 D_2). \tag{6.2}$$

Similarly, the quadratic risk of the RE of $\boldsymbol{\beta}$ is found to be

$$R_2^*(\hat{\boldsymbol{\beta}}; W_2) = E(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})' W_2 (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \sigma^2 tr(W_2 D_2^*) + \boldsymbol{\delta}' W_2 \boldsymbol{\delta}. \tag{6.3}$$

Then the quadratic risk of the SRE of $\boldsymbol{\beta}$ is obtained as

$$\begin{aligned} R_2(\hat{\boldsymbol{\beta}}_d; W_2) &= E[\hat{\boldsymbol{\beta}}_d - \boldsymbol{\beta}]' W_2 [\hat{\boldsymbol{\beta}}_d - \boldsymbol{\beta}] \\ &= d^2\sigma^2 tr(W_2 D_3) + (1-2d)\sigma^2 tr(W_2 D_2^*) + (1-d)^2 \boldsymbol{\delta}' W_2 \boldsymbol{\delta}. \end{aligned} \tag{6.4}$$

The quadratic risk of the SRE becomes that of the RE when $d = 0$.

Now, for the PTE of $\boldsymbol{\beta}$, the quadratic risk expression is given by

$$\begin{aligned} R_3^*(\hat{\boldsymbol{\beta}}^{pt}; W_2) &= E(\hat{\boldsymbol{\beta}}^{pt} - \boldsymbol{\beta})' W_2 (\hat{\boldsymbol{\beta}}^{pt} - \boldsymbol{\beta}) = \sigma^2 \Big\{ tr(W_2 D_2) - tr(W_2 D_3) G_{3,m}(l_\alpha; \Delta) \Big\} \\ &\quad + \boldsymbol{\delta}' W_2 \boldsymbol{\delta} \Big\{ 2G_{3,m}(l_\alpha; \Delta) - G_{5,m}(l_\alpha^*; \Delta) \Big\}. \end{aligned} \tag{6.5}$$

Finally, the quadratic risk expression of the SPTE of $\boldsymbol{\beta}$ is given by

$$\begin{aligned} R_3(\hat{\boldsymbol{\beta}}_d^{pt}; W_2) &= E[\hat{\boldsymbol{\beta}}_d^{pt} - \boldsymbol{\beta}]' W_2 [\hat{\boldsymbol{\beta}}_d^{pt} - \boldsymbol{\beta}] \\ &= \sigma^2 \Big\{ tr(W_2 D_2) - (1-d^2) tr(W_2 D_3) G_{3,m}(l_\alpha; \Delta) \Big\} \\ &\quad + \boldsymbol{\delta}' W_2 \boldsymbol{\delta} \Big\{ 2(1-d) G_{3,m}(l_\alpha; \Delta) - (1-d)^2 G_{5,m}(l_\alpha^*; \Delta) \Big\}. \end{aligned} \tag{6.6}$$

The risk of the SPTE becomes that of the PTE when $d = 0$. The derivation of the above results is straightforward by using the Appendix B1 of Judge and Bock (1978).

6.1 Risk analysis for estimators

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 the analytical analyses of the quadratic risk function of the estimators of the slope parameter vector.

Comparison of UE and SRE

First consider the difference between the quadratic risks of the UE and SRE,

$$\begin{aligned} N_{12}(\tilde{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}_d; W_2) &= R_1(\tilde{\boldsymbol{\beta}}; W_2) - R_2(\hat{\boldsymbol{\beta}}_d; W_2) \\ &= \sigma^2 \{ tr(W_2 D_2) - d^2 tr(W_2 D_3) - (1-2d) tr(W_2 D_2^*) \} \\ &\quad - (1-d)^2 \boldsymbol{\delta}' W_2 \boldsymbol{\delta}. \end{aligned} \tag{6.7}$$

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

$$\frac{(1+d)}{(1-d)} tr\{W_2 D_3\} - \frac{2}{1-d} tr\{W D_2^*\} \gtreqless \frac{\boldsymbol{\delta}' W_2 \boldsymbol{\delta}}{\sigma^2}. \tag{6.8}$$

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

$$\lambda_1 \le \Big[\frac{\boldsymbol{\delta}' W_2 \boldsymbol{\delta}}{\boldsymbol{\delta}' D_3^{-1} \boldsymbol{\delta}} \Big] \le \lambda_2 \tag{6.9}$$

where λ_1 is the smallest and λ_2 is the largest characteristic roots of the matrix $[W_2D_3]$. Then we have $\Delta\lambda_1 \leq \left[\frac{\boldsymbol{\delta}' W_2 \boldsymbol{\delta}}{\sigma^2}\right] \leq \Delta\lambda_2$. Thus the risk of the SRE is bounded in the following way

$$R_1(\tilde{\boldsymbol{\beta}}; W_2) + \Delta\lambda_1 - tr(W_2D_2^*) \leq R_2(\hat{\boldsymbol{\beta}}_d; W_2) \leq R_1(\tilde{\boldsymbol{\beta}}; W_2) + \Delta\lambda_2 - tr(W_2D_2^*). \quad (6.10)$$

Clearly, when H_0 is true then $\Delta = 0$ and the bounds are equal. In a special case, if $W_2 = D_3^{-1}$ we get $\frac{tr(W_2D_3)}{\lambda_2} = \frac{tr(W_2D_3)}{\lambda_1} = 2$ and the difference between the risks becomes

$$N_{12}(\tilde{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}; W_2) \gtreqless 0 \quad \text{according as} \quad \Delta \lesseqgtr (1+d) - \psi_2 \quad (6.11)$$

where ψ_2 is defined in (6.27). In another special case, if $W_2 = I_2$ then the RE is superior to the UE if $\Delta \leq \frac{tr(W_2D_3)}{\lambda_2}$, which depends on the value of the elements of the matrix D_3.

Comparison of UE and SPTE

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

$$\begin{aligned} N_{13}(\tilde{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}_d^{pt}; W_2) &= R_1(\tilde{\boldsymbol{\beta}}; W_2) - R_3(\hat{\boldsymbol{\beta}}_d^{pt}; W_2) \\ &= (1-d^2)\sigma^2 tr(W_2D_3)G_{3,m}(l_\alpha; \Delta) - \boldsymbol{\delta}' W_2 \boldsymbol{\delta} \\ &\quad \times \Big\{2(1-d)G_{3,m}(l_\alpha; \Delta) - (1-d^2)G_{5,m}(l_\alpha^*; \Delta)\Big\}. \end{aligned} \quad (6.12)$$

Thus we have

$$\begin{aligned} N_{13}(\tilde{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}_d^{pt}; W_2) &\gtreqless 0 \quad \text{whenever} \\ \frac{\boldsymbol{\delta}' W_2 \boldsymbol{\delta}}{\sigma^2} &\lesseqgtr \frac{(1+d)tr(W_2D_3)G_{3,m}(l_\alpha; \Delta)}{\Big\{2G_{3,m}(l_\alpha; \Delta) - (1+d)G_{5,m}(l_\alpha^*; \Delta)\Big\}}. \end{aligned} \quad (6.13)$$

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

$$R_3^L \leq R_3(\hat{\boldsymbol{\beta}}_d^{pt}; W_2) \leq R_3^U \quad (6.14)$$

where

$$\begin{aligned} R_3^L &= R_1(\tilde{\boldsymbol{\beta}}; W_2) + \Delta\lambda_1\Big\{2G_{3,m}(l_\alpha; \Delta) - (1+d)G_{5,m}(l_\alpha^*; \Delta)\Big\} \\ R_3^U &= R_1(\tilde{\boldsymbol{\beta}}; W_2) + \Delta\lambda_2\Big\{2G_{3,m}(l_\alpha; \Delta) - (1+d)G_{5,m}(l_\alpha^*; \Delta)\Big\}. \end{aligned} \quad (6.15)$$

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

$$\begin{aligned} N_{13}(\tilde{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}_d^{pt}; W_2) &\leq 0 \quad \text{if} \quad \Delta \leq \frac{(1+d)tr(W_2D_3)G_{3,m}(l_\alpha; \Delta)}{\lambda_1\Big\{2G_{3,m}(l_\alpha; \Delta) - (1+d)G_{5,m}(l_\alpha^*; \Delta)\Big\}} \\ N_{13}(\tilde{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}_d^{pt}; W_2) &\geq 0 \quad \text{if} \quad \Delta \geq \frac{(1+d)tr(W_2D_2)G_{3,m}(l_\alpha; \Delta)}{\lambda_2\Big\{2G_{3,m}(l_\alpha; \Delta) - (1+d)G_{5,m}(l_\alpha^*; \Delta)\Big\}}. \end{aligned} \quad (6.16)$$

In a special case, if $W_2 = D_3^{-1}$ the difference between the risks becomes,

$$N_{13}(\tilde{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}_d^{pt}; W_2) \gtreqless 0 \quad \text{according as } \Delta \lesseqgtr \frac{2(1+d)G_{3,m}(l_\alpha;\Delta)}{\Big\{2G_{3,m}(l_\alpha;\Delta) - (1+d)G_{5,m}(l_\alpha^*;\Delta)\Big\}}. \tag{6.17}$$

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

$$R_3^0(\hat{\boldsymbol{\beta}}_d^{pt}; W_2) = \sigma^2\Big\{tr(W_2D_2) - (1-d^2)tr(W_2D_3)G_{3,m}(l_\alpha;0)\Big\} \tag{6.18}$$

which is equal to the risk of the UE for $d = 1$, but less than that of the UE whenever $d \neq 1$. However, as Δ moves away from 0, the risk of the SPTE increases and reaches the maximum at Δ_α (say) after crossing the risk line of the UE at $\Delta_{0\alpha}$ given by (6.18) then decreases towards $\sigma^2 tr(W_2D_2)$, the risk of the UE as $\Delta \to \infty$.

Comparison of SPTE and SRE

The difference between the quadratic risks of the SPTE and SRE is

$$\begin{aligned} N_{32}(\hat{\boldsymbol{\beta}}_d^{pt}, \hat{\boldsymbol{\beta}}_d; W_2) &= R_3(\hat{\boldsymbol{\beta}}_d^{pt}; W_2) - R_2(\hat{\boldsymbol{\beta}}_d; W_2) \\ &= \sigma^2\Big\{d^2tr(W_2D_3) + (1-2d)tr(W_2D_2^*) - tr(W_2D_2) \\ &\quad +(1-d^2)tr(W_2D_3)G_{3,m}(l_\alpha;\Delta)\Big\} - \boldsymbol{\delta}'W_2\boldsymbol{\delta} \\ &\quad \times\Big\{(1-d)2G_{3,m}(l_\alpha;\Delta) - (1-d^2)G_{5,m}(l_\alpha^*;\Delta) - (1-d)^2\Big\}. \end{aligned} \tag{6.19}$$

Thus we get

$$N_{32}(\hat{\boldsymbol{\beta}}_d^{pt}, \hat{\boldsymbol{\beta}}_d; W_2) \lesseqgtr 0 \text{ according as } \frac{\boldsymbol{\delta}'W_2\boldsymbol{\delta}}{\sigma^2} \gtreqless \frac{\Lambda^U(W_2, D, \Delta)}{\Lambda^L(G, d, \Delta)} \tag{6.20}$$

where

$$\Lambda^U(W_2, D, \Delta) = \Big\{d^2tr(W_2D_3) + (1-2d)tr(W_2D_2^*) - tr(W_2D_2) + (1-d^2)tr(W_2D_3)G_{3,m}(l_\alpha;\Delta)\Big\} \tag{6.21}$$

$$\Lambda^L(G, d, \Delta) = \Big\{(1-d)2G_{3,m}(l_\alpha;\Delta) - (1-d^2)G_{5,m}(l_\alpha^*;\Delta) - (1-d)^2\Big\}. \tag{6.22}$$

Therefore,

$$\begin{aligned} N_{32}(\hat{\boldsymbol{\beta}}_d^{pt}, \hat{\boldsymbol{\beta}}_d; W_2) &\geq 0 \text{ if } \Delta \leq \frac{\Lambda^U(W_2, D, \Delta)}{\lambda_1 \times \Lambda^L(G, d, \Delta)} \text{ and} \\ N_{32}(\hat{\boldsymbol{\beta}}_d^{pt}, \hat{\boldsymbol{\beta}}_d; W_2) &\leq 0 \text{ if } \Delta \geq \frac{\Lambda^U(W_2, D, \Delta)}{\lambda_2 \times \Lambda^L(G, d, \Delta)}. \end{aligned} \tag{6.23}$$

Under the H_0, $\Delta = 0$ and hence the risk-difference reduces to

$$\begin{aligned} N_{32}(\hat{\boldsymbol{\beta}}^{pt}, \hat{\boldsymbol{\beta}}, W_2, \Delta = 0) &= \sigma^2\Big\{tr(W_2D_3)\Big[d^2(1 - G_{3,m}(l_\alpha;0)) + G_{3,m}(l_\alpha;0)\Big] \\ &\quad + tr(W_2D_2^*)(1-2d) - tr(W_2D_2)\Big\}. \end{aligned} \tag{6.24}$$

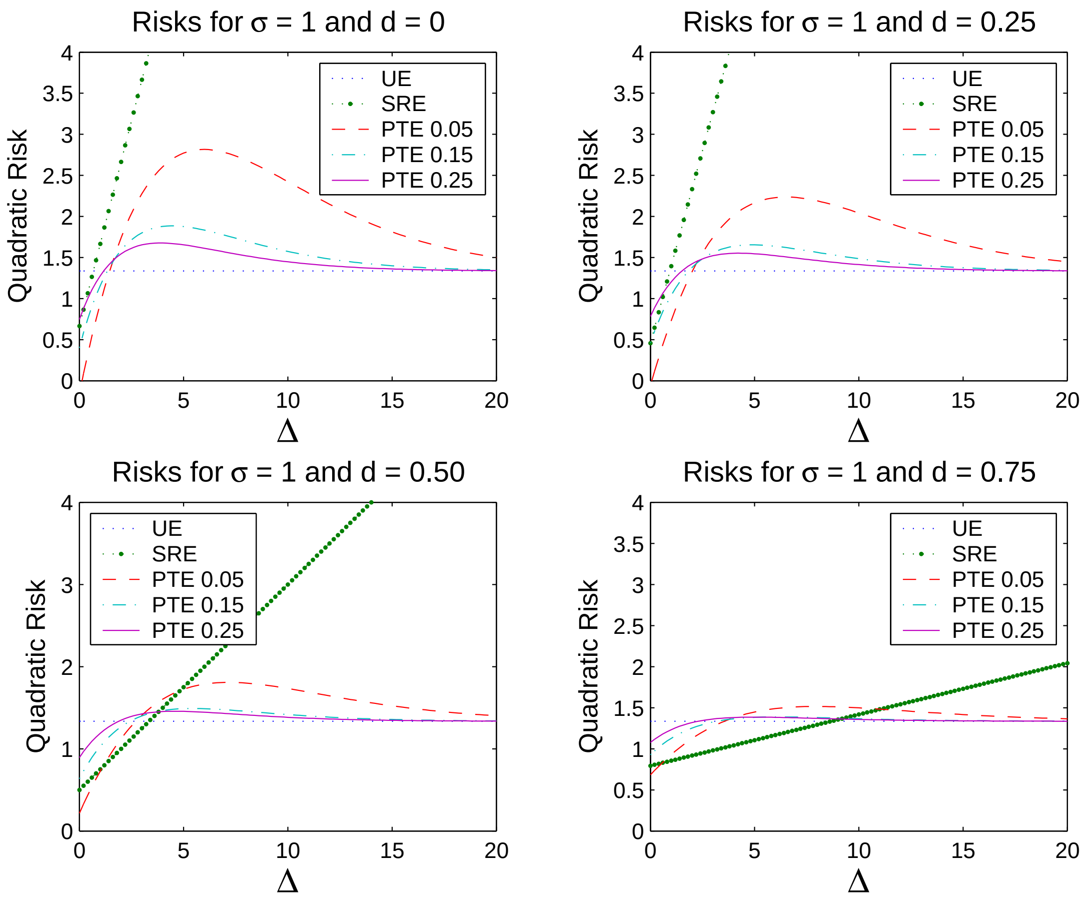

Figure 2: **Graph of Quadratic Risk functions of different estimators for selected values of d and $\sigma = 1$. Blood Pressure and Smoking Example.**

In a special case, when $W_2 = D_3^{-1}$,

$$N_{32}(\hat{\boldsymbol{\beta}}^{pt}, \hat{\boldsymbol{\beta}}, W_2) \quad \gtreqless \quad 0 \quad \text{according as}$$

$$\Delta \quad \lesseqgtr \quad \frac{2[d^2(1 - G_{3,m}(l_\alpha; \Delta)) + G_{3,m}(l_\alpha; \Delta)] + (1 - 2d)\psi_2 - \psi_1}{2(1 - d)G_{3,m}(l_\alpha; \Delta) - (1 - d^2)G_{5,m}(l_\alpha^*; \Delta) - (1 - d)^2} \tag{6.25}$$

where

$$\psi_1 \quad = \quad tr(D_3^{-1}D_2) = \frac{3(nQ)^2}{(n_1Q_1 + nQ)(n_2Q_2 + nQ)} \tag{6.26}$$

$$\psi_2 \quad = \quad tr(D_3^{-1}D_2^*) = \frac{n_1Q_1(n_2Q_2 + nQ) + n_2Q_2(n_1Q_1 + nQ)}{(n_1Q_1 + nQ)(n_2Q_2 + nQ)}. \tag{6.27}$$

6.2 Graphical Analysis of Quadratic Risks

Figures 2 displays the graphs of the quadratic risk function of the estimators against the non-centrality parameter for some selected values of d and $\sigma = 1$. The risk of the UE is constant and hence remains the same for all values of Δ. However, this constant risk increases with the increase in the value of σ. The quadratic risk of the SRE is unbounded and increases as the value of Δ grows large for all values of d, except $d = 1$. 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 SRE performs the worst. Like the quadratic bias, the rate of growth of the quadratic risk of the SRE declines as the value of d moves away from 0 to 1. However, at $d = 1$ the quadratic risk of the SRE is a constant.

The quadratic risk function of the PTE depends on the selected level of significance and the coefficient of distrust. When the null hypothesis is true, the SPTE has the smallest risk among the three estimators, regardless of the value of α. However, the minimum risk increases as the value of d grows larger. The difference among the different values of the quadratic risk of the SPTE for various values of α reduces as the value of d approaches to 1. This domination of the SPTE over the UE and SRE continues up to some small value of Δ, (say Δ_P), and than the risk function of the SPTE crosses that of the UE from the below and slowly grows up to a maximum for some moderate value of Δ. Then it declines gradually towards the risk curve of the UE. All the graphs in Figures 2 and 3 show the same behaviour of the SPTE 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, d and α. 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 SPTE is better than the UE if Δ is small and $d = 0$ or near 0. For moderate values of Δ, the SPTE is worse than the UE. This is more so when α is small.

The shape of the graphs of the quadratic risk functions and their properties do not depend on the choice of any particular data set.

7 An Example from Health Study

To demonstrate the application of the method, we consider a data set on a health study from Plank (2001, p.8.27). The study investigates the systolic blood pressure of a group of patients divided in to the smoking and non-smoking categories. In the sample there are 10 smokers and 11 non-smokers. The age of the patients is the explanatory variable, X, and is divided in to X_1, the age of the smoking patients and X_2, that of the non-smoking patients. The systolic blood pressure is the response variable, Y. Regression lines of Y on X_1 and Y on X_2 have been fitted to the data for the two group of patients separately. The scatterplot and the fitted regression lines are given in Figure 3. The fitted regression lines for the two groups of data are

$$\hat{y}_1 = -21.9487 + 3.0911x_1, \ (R_1^2 = 0.9512) \quad (7.1)$$

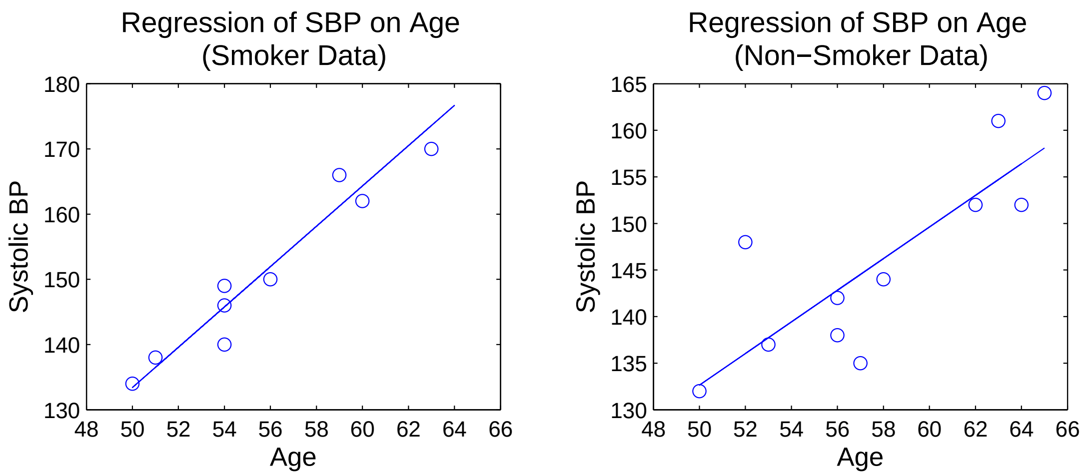

Figure 3: **Graph of the Least Squares Regression Line of Systolic BP on Age by Smoking Habit.**

$$\hat{y}_2 \quad = \quad 47.7437 + 1.6978x_2, \; (R_2^2 = 0.6761). \tag{7.2}$$

Other statistics relevant to the current study are $n_1Q_1 = 208.5$, $n_2Q_2 = 259.64$, $nQ = 468.14$ and $\hat{\beta} = 2.3184$. The observed value of the test statistic is 5.555 with a P-value of 0.0307. 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.

8 Concluding remarks

In this paper we have defined the unrestricted, shrinkage restricted and shrinkage preliminary test estimators for the slope parameters of the two suspected parallel regression models. The performances of the three different estimators of the slope parameter vector have been analyzed by using the criteria of quadratic bias and risk under quadratic loss. The SPTE has always smaller quadratic bias than the SRE for values of $d \neq 1$, 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 6 and the graphs displayed in Figures 2. The SRE is the best if and only if $\Delta = 0$. In the face of uncertainty on the value of Δ, if Δ is likely to be small, then the SPTE is the preferred estimator, 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 SPTE reaches its maximum. For very large values of Δ the SPTE performs as good as the UE under the quadratic risk criterion, but a lot better than the SRE.

In practice, the prior information on the equality of slopes would either come from expert knowledge of the data generating process or from the results of previous studies. In either case, the value of Δ is very unlikely to be too large, but most likely to be close to 0. Also, in such a situation the value of d would be closer to 0, rather than 1. If d is close to 0, the reliability of the prior information is too low. Thus in most realistic situation, Δ is likely to be close to 0 and d should not be far away from 0. Therefore, under the above situation the SPTE would be the best choice to guarantee the minimum quadratic risk.

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 comments by two unknown referees that help improve the presentation 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 coefficient . *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.

James, W. and Stein, C. (1961): Estimation with quadratic loss. *Proceedings of the Fourth Berkeley Symposium on Math. Statist. and Probability*, University of California Press, Berkeley, 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. (2003). Estimation of the Parameters of two Parallel Regression Lines Under Uncertain Prior Information, *Biometrical Journal*, Vol. 44, 73-90.

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

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 predictions: Pooling two identical regression lines. *Jou. Appld. Statist. Sc.*, **7**, 63-86.

Plank, A.W. (2001). Experimental Design, Distance Education Centre, University of Southern Queensland, Australia.

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

Saleh, A.K.Md.E. (2006). Theory of Preliminary Test and Stein-Type Estimation with Applications, 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.

Stein, C. (1956): Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proceedings of the Third Berkeley Symposium on Math. Statist. and Probability*, University of California Press, Berkeley, Vol. 1, 197–206.