Test for intercept after pre-testing on slope - a robust method

Rossita M Yunus* and Shahjahan Khan
Department of Mathematics and Computing
Australian Centre for Sustainable Catchments
University of Southern Queensland
Toowoomba, Q 4350, AUSTRALIA.
Emails:yunus@usq.edu.au and khans@usq.edu.au

Abstract

The idea of using non-sample prior information in the form of pre-testing for improving properties of estimators is applied in the testing regime to achieve better power of the ultimate test in this paper. In particular, to test the intercept of a simple regression model, prior information from previous investigations or expert knowledge on the suspected value of the slope is potentially beneficial. Any uncertainty on the value of the slope is removed by performing a pre-test before testing the significance of the intercept. The impact of the pre-test on the performance (power and size) of the ultimate test is studied. A robust procedure based on M-estimator is used to formulate a test and deriving its power function. It is shown that the ultimate test based pre-test achieves a reasonable dominance over the others asymptotically and performs better for larger coefficient of variation.
Keywords: pre-test, asymptotic size, asymptotic power, M-estimation, regression model.

1 Introduction

Consider a simple regression model of n observable random variables, $X_i,\ i = 1, \ldots, n$

$$X_i = \theta + \beta c_i + e_i, \tag{1.1}$$

where the errors e_i's are from an unspecified symmetric and continuous distribution function, $F_i,\ i = 1, \ldots, n$, the c_i's are known real constants of the explanatory variable and θ and β are the unknown intercept and slope parameters respectively.

Testing the significance of the intercept, $H_0^\star : \theta = 0$ are carried out under three possible cases based on the knowledge of the slope. In the first case, the slope is unspecified i.e. it is treated as a nuisance parameter and the testing of the significance of the intercept is referred as the unrestricted test (UT). If a non-sample prior information on the value of the slope (say

*on leave from Institute of Mathematical Sciences, Faculty of Sciences, University of Malaya, Malaysia.

0) is known, the testing of the significance of the intercept is defined as the restricted test (RT). If the prior information on the value of the slope is uncertain, it is suggested to perform a pre-test to remove the uncertainty on the value of the slope before testing the significance of the intercept. For the final case, the ultimate test (on the intercept) is defined as the pre-test test (PTT). Obviously the pre-test (on the slope) affects the power and size of the ultimate test (on the intercept).

In spite of numerous works in improving the estimation by the inclusion of a non-sample prior information (Saleh, 2006, Khan and Saleh, 2001 and Khan et al., 2002), very little attention has been paid in improving the performance of the test of the parameters. The effect of pre-test on the performance (size and power) of the ultimate test are studied in parametric cases (Bechhofer, 1951 and Bozivich et al., 1956) as well as non-parametric cases (Saleh and Sen, 1982) though the number of studies is very small in literature. Saleh and Sen (1981) use rank based nonparametric tests and formulate the power function of the ultimate test. However, there are some limited discussions in investigating the power of the ultimate test discussed in the paper.

Realizing M-estimation is more popular than the other robust methods and well known for its flexibility and well defined for a variety of models for which MLE is also defined (Huber, 1981, p.43, Jurěcková and Sen, 1996 p.80), a score type M-test is proposed to formulate the power functions of the UT, RT and PTT (see Yunus and Khan, 2007). The asymptotic power functions for the UT, RT and PTT that are derived using M-test are found to have the same form as that derived by using the rank statistic by Saleh and Sen (1982) though the methodology of M-estimation and R-estimation is different. The paper discussed the asymptotic comparison of the UT, RT and PTT analytically and computationally for a special case of the value of coefficient of variation. The dependency of power functions to the coefficient of variation is studied in this paper as an extension of the previous works.

Along with some preliminary notions, the method of M-estimation is presented and statistical tests concerning testing on the intercept, namely, the UT, RT and PTT are given in Section 2. The asymptotic distributions of the test statistics and the asymptotic power functions of the test are given in Section 3. Section 4 is devoted to the analytical results comparing the asymptotic power functions of the UT, RT and PTT while the investigation of the power functions through an illustrative example is presented in Section 5.

2 The Proposed Test

Given an absolutely continuous function $\rho : \Re \rightarrow \Re$, M-estimator of θ and β is defined as the values of θ and β that minimize the objective function

$$\sum_{i=1}^{n} \rho(X_i - \theta - \beta c_i). \tag{2.1}$$

M-estimator of θ and β can also be defined as the solutions of the system of equations,

$$\begin{array}{rcccl} \sum_{i=1}^{n} \psi_\theta(X_i) & = & \sum_{i=1}^{n} \psi(X_i - \theta - \beta c_i) & = & 0, \\ \sum_{i=1}^{n} \psi_\beta(X_i) & = & \sum_{i=1}^{n} c_i \psi(X_i - \theta - \beta c_i) & = & 0. \end{array} \tag{2.2}$$

If ρ is differentiable with partial derivatives $\psi_\theta = \partial\rho/\partial\theta$ and $\psi_\beta = \partial\rho/\partial\beta$, then the M-estimators that minimize the function in (2.1) are the solutions to the system (2.2). On the contrary, the M-estimators obtained from solving system (2.2) may not minimize equation (2.1) (c.f. Caroll and Rupert, 1988 p.210). The system of equations (2.2) may have more roots, while only one of them leads to a global minimum of (2.1). Jurěcková and Sen (1996) have given proof that there exists at least one root of (2.2) which is a $\sqrt{n}$ - consistent estimator of θ and β under some conditions [c.f. p.215 - 224]. The ψ function is decomposed into the sum

$$\psi = \psi_a + \psi_c + \psi_s,$$

where

(a) ψ_a is absolutely continuous function with absolutely continuous derivative.

(b) ψ_c is a continuous, piecewise linear function with knots at $\mu_1, \ldots, \mu_k$, that is, constant in a neighborhood of $\pm\infty$ and hence its derivative is a step function $\psi_c'(z) = \alpha_v,\ \mu_v < z < \mu_{v+1},\ v = 0, 1, \ldots, k$ where $\alpha_0, \ldots, \alpha_k \in \Re,\ \alpha_0 = \alpha_k = 0$ and $-\infty = \mu_0 < \mu_1 < \ldots < \mu_{k+1} = \infty$. We assume that $f(z) = \frac{dF(z)}{dz}$ is bounded in neighborhoods of $S_{\mu_1}, \ldots, S_{\mu_k}$.

(c) ψ_s is a nondecreasing step function, $\psi_s(z) = \lambda_v,\ q_v < z \leq q_{v+1}, v = 1, \ldots, m$ where $-\infty = q_0 < q_1 < \ldots < q_{m+1} = \infty$ and $-\infty < \lambda_0 < \lambda_1 < \ldots < \lambda_m < \infty$. We assume that $0 < f(z) = (d/dz)F(z)$ and $f'(z) = (d^2/dz^2)F(z)$ are bounded in neighborhoods of $S_{q_1}, \ldots, S_{q_m}$.

The asymptotic result under conditions M1 to M5 of Jurěcková and Sen (1996, p.217) is used in this paper. Further assume that all ψ_a, ψ_c and ψ_s are nondecreasing and skew symmetric that is $\psi_j(-x) = -\psi_j(x),\ j = 1, 2, 3$. Let F be symmetric about 0, so that

$$\int_{-\infty}^{\infty} \psi(x) dF(x) = 0.$$

Assume that

$$\sigma_0^2 = \int_{-\infty}^{\infty} \psi^2(x) dF(x). \tag{2.3}$$

Following Jurěcková and Sen (1996, p.217), two cases are considered:

(i) if $\psi_s = 0$ then

$$\gamma = \int_{-\infty}^{\infty} (\psi_a'(x) + \psi_c'(x)) f(x) dx. \tag{2.4}$$

(ii) if $\psi_a = \psi_c = 0$, then

$$\gamma = \sum (\lambda_v - \lambda_{v-1}) f(S_{q_v}). \tag{2.5}$$

Further assume that σ_0 and γ are both positive and finite quantities. Let the distribution function, F be continuous and symmetric about zero and have finite Fisher information,

$$I(f) = \int_{-\infty}^{\infty} \{f'(x)/f(x)\}^2 dF(x), \tag{2.6}$$

where $f'(x) = (d/dx)f(x) = (d^2/dx^2)F(x)$. Assume that

(i) there exists finite constants $\bar{c}$ and $C^\star(> 0)$ such that

$$\lim_{n\to\infty} \bar{c}_n = \bar{c} \quad \text{and} \quad \lim_{n\to\infty} n^{-1}C_n^{\star 2} = C^{\star 2} \tag{2.7}$$

with

$$\bar{c}_n = n^{-1}\sum_{i=1}^{n} c_i \quad \text{and} \quad C_n^{\star 2} = \sum_{i=1}^{n} c_i^2 - n\bar{c}_n^2 \tag{2.8}$$

both exist.

(ii) the c_i's are all bounded, so that by (i),

$$\max_{1\le i\le n} (c_i - \bar{c}_n)^2/C_n^{\star 2} \to 0, \quad \text{as } n \to \infty. \tag{2.9}$$

Let $\psi : \Re \to \Re$ be nondecreasing and skew symmetric score function. For any real numbers a and b, consider the statistics below

$$M_{n_1}(a, b) = \sum_{i=1}^{n} \psi(X_i - a - bc_i), \tag{2.10}$$

$$M_{n_2}(a, b) = \sum_{i=1}^{n} c_i\psi(X_i - a - bc_i). \tag{2.11}$$

If β is unspecified, the designated test function is ϕ_n^{UT} with the null hypothesis $H_0^\star : \theta = 0$. The testing for θ involves the elimination of the nuisance parameter β. It follows that $M_{n_2}(0, b)$ is decreasing if b is increasing (Jurěcková and Sen, 1996, p.85) and under local hypothesis, $H_0^{(1)} : \beta = 0$, $M_{n_2}(0, 0)$ has expectation 0. Then let

$$\tilde{\beta} = (\sup\{b : M_{n_2}(0, b) > 0\} + \inf\{b : M_{n_2}(0, b) < 0\})/2.$$

Then $\tilde{\beta}$ is a translation invariant and robust estimator of β.

We consider the test statistic $T_n^{UT} = M_{n_1}(0, \tilde{\beta})$ where under $H_0^\star$, as $n \to \infty$,

$$\frac{T_n^{UT}}{\sqrt{C_n^{(1)} S_n^{(1)2}}} \to N(0, 1) \tag{2.12}$$

with $C_n^{(1)} = n - n^2\bar{c}_n^2/\sum c_i^2 = nC_n^{\star 2}/(C_n^{\star 2} + n\bar{c}_n^2)$ and $[S_n^{(1)}]^2 = \sum \psi^2(x_i - \tilde{\beta}c_i)/n$.

If $\beta = 0$, the designated test function is ϕ_n^{RT} for testing the null hypothesis $H_0^\star : \theta = 0$. The proposed test statistic is $T_n^{RT} = M_{n_1}(0, 0)$. Note that for large n, under $H_0 : \theta = 0, \beta = 0$,

$$n^{-1/2}T_n^{RT} = n^{-1/2}M_{n_1}(0, 0) \to N(0, \sigma_0^2), \tag{2.13}$$

where $\sigma_0^2 = \int_{-\infty}^{\infty} \psi^2(x)dF(x)$ (see Sen, 1982, eq 3.7).

For the preliminary test on the slope, the test function, ϕ_n^{PT} is designed to test the null hypothesis $H_0^{(1)} : \beta = 0$. The proposed test statistic is $T_n^{PT} = M_{n_2}(\tilde{\theta}, 0)$ where

$$\tilde{\theta} = (\sup\{a : M_{n_1}(a, 0) > 0\} + \inf\{a : M_{n_1}(a, 0) < 0\})/2$$

is a robust estimator. Under $H_0^{(1)}$, as $n \to \infty$,

$$\frac{T_n^{PT}}{\sqrt{C_n^{(3)} {S_n^{(3)}}^2}} \to N(0, 1), \tag{2.14}$$

where $C_n^{(3)} = \sum c_i^2 - n\bar{c}_n^2 = C_n^{\star 2}$ and $[S_n^{(3)}]^2 = \sum \psi^2(x_i - \tilde{\theta})/n$.

The consistency of $[S_n^{(1)}]^2$, $[S_n^{(2)}]^2 = \sum \psi^2(x)/n$ and $[S_n^{(3)}]^2$ as estimators of σ_0^2 follows by law of large number (Jurečková and Sen, 1981).

Now, we are in a position to formulate a test function ϕ_n^{PTT} to test $H_0^\star : \theta = 0$ following a preliminary test on β. First, we consider the case where all of $\phi_n^{(j)}$, $j = 1, 2, 3$ are one-sided test. Let us choose positive numbers α_j $(0 < \alpha_j < 1)$ and real values $\ell_{n,\alpha_j}^{(j)}$, $j = 1, 2, 3$, such that for large sample size,

$$P[T_n^{UT} > \ell_{n,\alpha_1}^{UT} | H_0^\star : \theta = 0] = \alpha_1, \tag{2.15}$$

$$P[T_n^{RT} > \ell_{n,\alpha_2}^{RT} | H_0 : \theta = 0, \beta = 0] = \alpha_2, \tag{2.16}$$

$$P[T_n^{PT} > \ell_{n,\alpha_3}^{PT} | H_0^{(1)} : \beta = 0] = \alpha_3, \tag{2.17}$$

where $\ell_{n,\alpha_j}^{(j)}$ is the critical value of $T_n^{(j)}$ at the α_j level of significance. Let $\Phi(x)$ be the standard normal cumulative distribution function, then

$$\Phi(\tau_\alpha) = 1 - \alpha, \quad \text{for } \ 0 < \alpha < 1. \tag{2.18}$$

Using equations (2.12)–(2.18), as $n \to \infty$ we have

$$\frac{n^{-1/2}\ell_{n,\alpha_2}^{RT}}{\sqrt{{S_n^{(2)}}^2}} \to \tau_{\alpha_2} = \frac{n^{-1/2}\ell_{n,\alpha_2}^{RT}}{\sqrt{\sigma_0^2}} \ \text{(say)}. \tag{2.19}$$

$$\frac{n^{-1/2}\ell_{n,\alpha_1}^{UT}}{\sqrt{{S_n^{(1)}}^2 C_n^{(1)}/n}} \to \tau_{\alpha_1} = \frac{n^{-1/2}\ell_{n,\alpha_1}^{UT}}{\sqrt{\sigma_0^2 C^{\star 2}/(C^{\star 2} + \bar{c}^2)}} \ \text{(say)}, \tag{2.20}$$

$$\frac{n^{-1/2}\ell_{n,\alpha_3}^{PT}}{\sqrt{{S_n^{(3)}}^2 C_n^{\star 2}/n}} \to \tau_{\alpha_3} = \frac{n^{-1/2}\ell_{n,\alpha_3}^{PT}}{\sqrt{\sigma_0^2 C^{\star 2}}} \ \text{(say)}, \tag{2.21}$$

where ${S_n^{(1)}}^2 = \sum \psi^2(x_i - \tilde{\beta}c_i)/n$, $C_n^{(1)} = n - n^2\bar{c}_n^2/\sum c_i^2$, ${S_n^{(3)}}^2 = \sum \psi^2(x_i - \tilde{\theta})/n$, and $C_n^{\star 2} = \sum c_i^2 - n\bar{c}_n^2$.

Now we may write

$$\phi_n^{PTT} = I\left[(T_n^{PT} \le \ell_{n,\alpha_3}^{PT}, T_n^{RT} > \ell_{n,\alpha_2}^{RT}) \ \text{ or } \ (T_n^{PT} > \ell_{n,\alpha_3}^{PT}, T_n^{UT} > \ell_{n,\alpha_1}^{UT})\right] \tag{2.22}$$

as the test function for testing $H_0^\star : \theta = 0$ after a pre-test on β. Note that $I(A)$ stands for the indicator function of the set A. It takes value 1 if A occurs, otherwise 0. The function enables us to define the power of the test ϕ_n^{PTT}, that is given by

$$\begin{aligned} \Pi_n^{PTT}(\theta) &= E(\phi_n^{PTT}|\theta) \\ &= P[T_n^{PT} \leq \ell_{n,\alpha_3}^{PT}, T_n^{RT} > \ell_{n,\alpha_2}^{RT}|\theta] + P[T_n^{PT} > \ell_{n,\alpha_3}^{PT}, T_n^{UT} > \ell_{n,\alpha_1}^{UT}|\theta]. \end{aligned} \quad (2.23)$$

In general, the power of the test ϕ_n^{PTT} depends on $\alpha_1, \alpha_2, \alpha_3, \theta, n$ as well as β. Note that the size of the ultimate test α_n^{PTT} is a special case of the power of the test when $\theta = 0$. Since the nuisance parameter β is unknown, but, suspected to be close to 0, it is of interest to study the dependence of both α_n^{PTT} and $\Pi_n^{PTT}(\theta)$ on β (close to 0).

3 Asymptotic properties for UT, RT and PTT

Let $\{K_n\}$ be a sequence of alternative hypotheses, where

$$K_n : (\theta, \beta) = (n^{-1/2}\lambda_1, n^{-1/2}\lambda_2), \quad (3.1)$$

with λ_1, λ_2 are (fixed) real numbers. The following results follows directly from Yunus and Khan (2007) [and hence, their proofs are omitted]: Under $\{K_n\}$, for large sample,

- $$n^{-1/2}\begin{bmatrix} T_n^{UT} \\ T_n^{PT} \end{bmatrix} \sim N_2\left[\begin{pmatrix} \frac{\gamma\lambda_1 C^{\star 2}}{C^{\star 2}+\bar{c}^2} \\ \gamma\lambda_2 C^{\star 2} \end{pmatrix}, \sigma_0^2\begin{pmatrix} \frac{C^{\star 2}}{C^{\star 2}+\bar{c}^2} & -\frac{\bar{c}C^{\star 2}}{C^{\star 2}+\bar{c}^2} \\ -\frac{\bar{c}C^{\star 2}}{C^{\star 2}+\bar{c}^2} & C^{\star 2} \end{pmatrix}\right] \quad (3.2)$$

  and

- $$n^{-1/2}\begin{bmatrix} T_n^{RT} \\ T_n^{PT} \end{bmatrix} \sim N_2\left[\begin{pmatrix} \gamma(\lambda_1 + \lambda_2\bar{c}) \\ \gamma\lambda_2 C^{\star 2} \end{pmatrix}, \sigma_0^2\begin{pmatrix} 1 & 0 \\ 0 & C^{\star 2} \end{pmatrix}\right]. \quad (3.3)$$

Define $d(q_1, q_2 : \rho)$ to be the bivariate normal probability integral for random variables x and y,

$$d(q_1, q_2; \rho) = \frac{1}{2\pi(1-\rho^2)^{1/2}} \int_{q_1}^{\infty}\int_{q_2}^{\infty} \exp\left\{\frac{-(x^2 + y^2 - 2\rho xy)}{2(1-\rho^2)}\right\} dxdy, \quad (3.4)$$

where q_1, q_2 are real numbers and $-1 < \rho < 1$. Here $d(q_1, q_2; \rho)$ is the complement of the cumulative density function of standard bivariate normal variable.

Using equations (2.18)−(2.21), (3.2)−(3.4), the power function for the PTT, under $\{K_n\}$, is

$$\begin{aligned} &\Pi_n^{PTT}(\lambda_1, \lambda_2) = E(\phi_n^{PTT}|K_n) \to \Pi^{PTT}(\lambda_1, \lambda_2) \\ &= \Phi(\tau_{\alpha_3} - \gamma\lambda_2 C^\star/\sigma_0)[1 - \Phi(\tau_{\alpha_2} - \gamma(\lambda_1 + \lambda_2\bar{c})/\sigma_0)] \; + \\ &\quad d(\tau_{\alpha_3} - \gamma\lambda_2 C^\star/\sigma_0, \tau_{\alpha_1} - \gamma\lambda_1\sqrt{C^{\star 2}/(C^{\star 2} + \bar{c}^2)}/\sigma_0; -\bar{c}/\sqrt{C^{\star 2} + \bar{c}^2}\,). \end{aligned} \quad (3.5)$$

Similarly, the power function for the RT and UT are respectively

$$\Pi^{RT}(\lambda_1, \lambda_2) = 1 - \Phi(\tau_{\alpha_2} - \gamma(\lambda_1 + \lambda_2 \bar{c})/\sigma_0) \tag{3.6}$$

and

$$\Pi^{UT}(\lambda_1, \lambda_2) = 1 - \Phi(\tau_{\alpha_1} - \gamma\lambda_1 \sqrt{C^{\star 2}/(C^{\star 2} + \bar{c}^2)}\, /\sigma_0) \tag{3.7}$$

using equations (2.18)−(2.20).

4 Asymptotic comparison

The below results are obtained by Yunus and Khan (2007): For $\bar{c} > 0$, $\alpha_1 = \alpha_2 = \alpha$ and $\lambda_2 \geq 0$, it is easy to show that $\lambda_1 + \lambda_2 \bar{c} \geq \lambda_1 \sqrt{C^{\star 2}/(C^{\star 2} + \bar{c}^2)}$. Thus,

- $\Pi^{RT}(\lambda_1, \lambda_2) \geq \Pi^{PTT}(\lambda_1, \lambda_2)$
- $\Pi^{RT}(\lambda_1, \lambda_2) > \Pi^{UT}(\lambda_1, \lambda_2)$
- $\Pi^{UT}(\lambda_1, \lambda_2) - \Pi^{PTT}(\lambda_1, \lambda_2) \lesseqqgtr 0$ if $B \lesseqqgtr |A|$ where $A = \Phi(\tau_\alpha - \gamma(\lambda_1 + \lambda_2 \bar{c})/\sigma_0) - \Phi(\tau_\alpha - \gamma\lambda_1 \sqrt{C^{\star 2}/(C^{\star 2} + \bar{c}^2)}/\sigma_0)$ and $B = d(\tau_{\alpha_3} - \gamma\lambda_2 C^\star/\sigma_0, \tau_\alpha - \gamma(\lambda_1 + \lambda_2\bar{c})/\sigma_0; 0) - d(\tau_{\alpha_3} - \gamma\lambda_2 C^\star/\sigma_0, \tau_\alpha - \gamma\lambda_1\sqrt{C^{\star 2}/(C^{\star 2} + \bar{c}^2)}/\sigma_0; -\bar{c}/\sqrt{C^{\star 2} + \bar{c}^2}\,)$.

Let $\alpha_1 = \alpha_2 = \alpha$, $\bar{c} > 0$, $\lambda_2 \geq 0$ we find $\Pi^{RT}(0, \lambda_2) = 1 - \Phi(\tau_\alpha - \gamma\lambda_2\bar{c}/\sigma_0) \geq 1 - \Phi(\tau_\alpha) = \Pi^{UT}(0, \lambda_2)$ from equations (3.6) and (3.7). Obvious that if the inverse of the coefficient of variation that is $\bar{c}/\sigma_0$ decreases, $\Pi^{RT}(0, \lambda_2)$ decreases but $\Pi^{UT}(0, \lambda_2)$ fixed at α.

The size of the PTT is always smaller than that of the RT and hence it approaches the size of the UT (that is constant at α) as the coefficient of variation increases.

The analytical results in this section is accompanied with an illustrative example in investigating the comparison of the power of the tests discussed in the next section. The power of the tests at any value other than $\theta = 0$ is also considered in the example to study the behavior of the power functions corresponds to the probabilities of type I and type II errors.

5 Illustrative Example - Power Comparison

For this illustrative example, the random errors of the simple linear model are generated from Normal distribution with mean 0 and variance 1. The sample size is $n = 30$. Four sets of values: 0 and 1 with ratio 1:9, 3:7, 5:5, 7:3 are considered as the values of the regressor c_i, $i = 1, 2, \ldots, 30$. These values guarantee $\bar{c_n}^2/C_n^{\star 2}$ to be 0.300, 0.078, 0.033, 0.014. Note that as the coefficient of variation $(\sigma_0/\bar{c})$ increases, $\bar{c_n}^2/C_n^{\star 2}$ decreases.

In this example, the ψ function is taken as Huber ψ function (Hoaglin et al., 1983, p.366, Wilcox, 2005, p.77), is defined as

$$\psi_h(u_i) = \begin{cases} u_i & \text{if} \quad |u_i| \leq k \\ k\ sgn(u_i) & \text{if} \quad |u_i| > k, \end{cases}$$

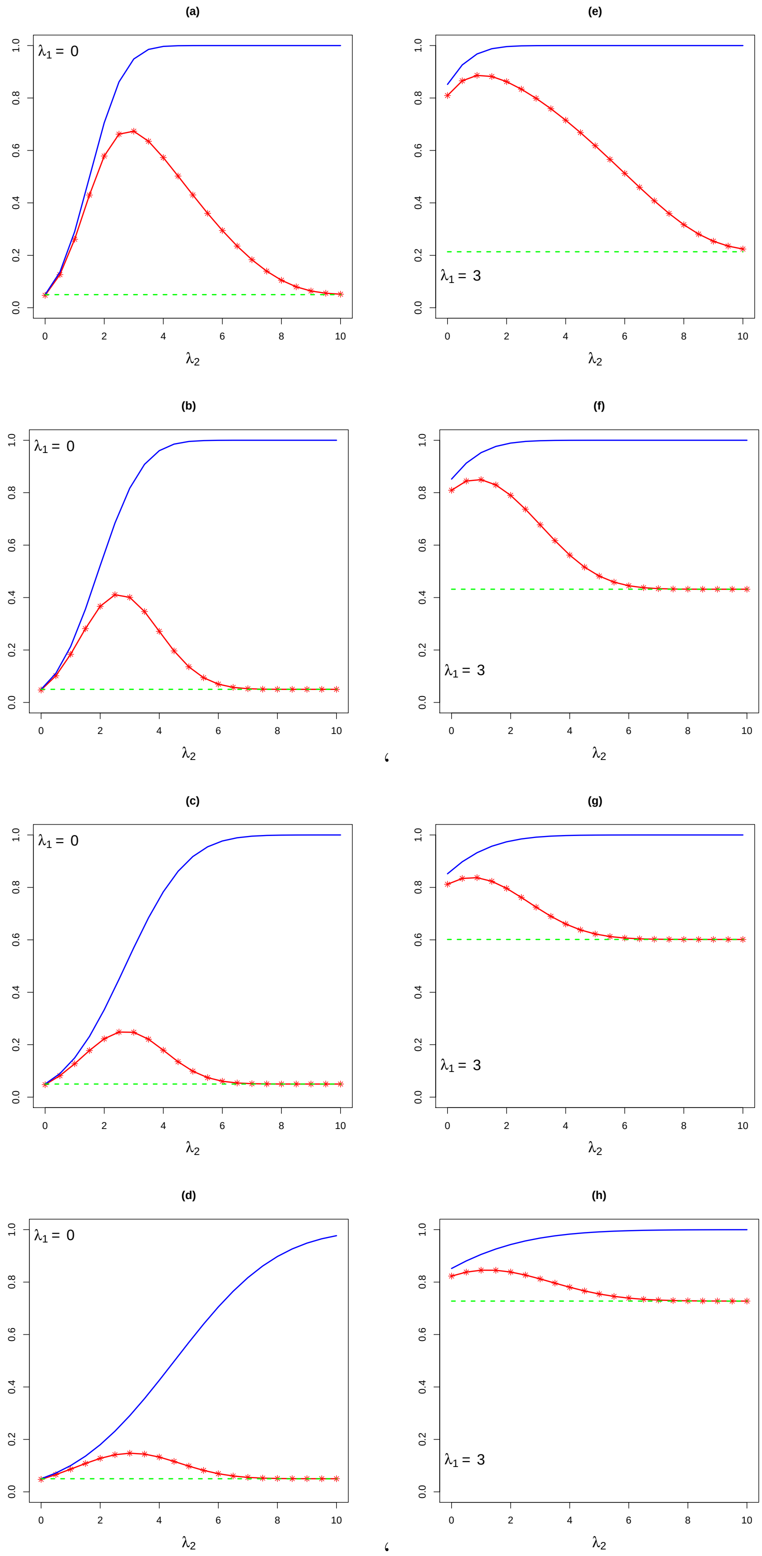

Figure 1: Graphs of power functions for increasing λ_2 with $\alpha_1 = \alpha_2 = \alpha_3 = \alpha = 0.05$ and $n = 30$ for selected values of λ_1. Dotted line, solid line and line with star represent $\Pi^{UT}(\lambda_1, \lambda_2)$, $\Pi^{RT}(\lambda_1, \lambda_2)$ and $\Pi^{PTT}(\lambda_1, \lambda_2)$ respectively. Here $\bar{c_n}^2/C_n^{\star 2} = 0.300$ for graphs (a) & (e), $\bar{c_n}^2/C_n^{\star 2} = 0.078$ for graphs (b) & (f), $\bar{c_n}^2/C_n^{\star 2} = 0.033$ for graphs (c) & (g) and $\bar{c_n}^2/C_n^{\star 2} = 0.014$ for graphs (d) & (h).

where $u_i = X_i - \theta - \beta c_i$. As suggested in many reference books (Wilcox, 2005, p.76), the value of $k = 1.28$ is chosen because $k = 1.28$ is the 0.9 quantile of a standard normal distribution, there is a 0.8 probability that a randomly sampled observations will have a value between $-k$ and k (Wilcox, 2005, p.76). The estimate for σ_0 is taken to be $\sum \psi(u)^2/n$. The estimate for γ is $\sum \psi'(u)/n$ (Caroll and Rupert, 1988, p.212) where

$$\psi'(u) = \begin{cases} 1 & \text{if} \quad |u| \le 1.28 \\ 0 & \text{if} \quad |u| > 1.28. \end{cases}$$

The Π^{UT}, Π^{RT} and Π^{PTT} are calculated using the formulas given by equations (3.5), (3.6) and (3.7). The R-package (mvtnorm) is used in computing the bivariate Normal probability integral.

The size of the UT, RT and PTT are plotted against λ_2 in Figures 1(a)$-$(d) while power of the test at $\lambda_1 = 3$ are plotted against λ_2 in Figures 1(e)$-$(h) for selected values of $\bar{c_n}^2/C_n^{\star 2}$. We desire the size of a particular test to be small and power of the test to be large so that probability of type I and II errors are both small. As λ_2 grows larger, the size and power of the RT grows larger and approaches 1 (see Figure 1(a) & (e)). Observe also the size of the RT approaches 1 faster for larger values of $\bar{c_n}^2/C_n^{\star 2}$ as λ_2 grows larger. From Figures 1(a)$-$(d), the size of the UT is constant at $\alpha = 0.05$ and does not depend on λ_2 and $\bar{c_n}^2/C_n^{\star 2}$. The power of the UT though constant regardless the values of λ_2, it increases as a smaller $\bar{c_n}^2/C_n^{\star 2}$ is selected (see Figures 1(e)$-$(h)). Starting at a value that is slightly less than α, the size of the PTT increases before drops and converges to the nominal size $\alpha = 0.05$ as λ_2 grows larger. The size of the PTT is small when $\bar{c_n}^2/C_n^{\star 2}$ is small and large for larger $\bar{c_n}^2/C_n^{\star 2}$ but it is always less than that of the RT. The power of the PTT is larger for smaller λ_2 and it decreases to the power of the UT as λ_2 grows larger. For a larger value of λ_2, the power of the PTT is increasing as $\bar{c_n}^2/C_n^{\star 2}$ is decreasing. The PTT has smaller size and larger power for a small $\bar{c_n}^2/C_n^{\star 2}$ than a larger one.

It is impossible to obtain a test that uniformly minimizes the size and maximizes the power at the same time. We are looking for a test that is a compromise between minimizing the size and maximizing the power (small probabilities of type I and type II errors). The RT is the best choice for its largest power but the worst choice for its largest size as λ_2 grows larger. On the contrary, the UT is the best choice for its smallest size but the worst choice for its smallest power. Both RT and UT uniformly minimize or maximize the size and power at the same time. The PTT has larger power than the UT for small and moderate values of λ_2 and it has significantly smaller size than that of the RT for moderate and large λ_2. Therefore, if our objective is to obtain a test that has better probabilities for both type I and type II errors, the PTT is suggested as the best option. The PTT is a compromise between minimizing the size and maximizing the power than the RT and UT. The PTT has a smaller size and a larger power for a larger coefficient of variation than a smaller one.

ACKNOWLEDGEMENTS

The authors thankfully acknowledge valuable suggestions of Professor A K Md E Saleh, Carleton University, Canada that helped improve the content and quality of the results in the paper.

References

Bechhofer, R.E. (1951). The effect of preliminary test of significance on the size and power of certain tests of univariate linear hypotheses. Ph.D. Thesis (unpublished), Columbia Univ.

Bozivich, H., Bancroft, T.A. and Hartley, H. (1956). Power of analysis of variance test procedures for certain incompletely specified models. *Ann. Math. Statist.* **27**, 1017 - 1043.

Caroll, R.J. and Rupert, D. (1988). *Transformation and Weighting in Regression.* Chapman & Hall, US.

Hoaglin, D.C., Mosteller, F. and Tukey, J.W. (1983). *Understanding Robust and Explanatory Data Analysis.* John Wiley and Sons, US.

Huber, P.J. (1981). *Robust Statistics.* Wiley, New York.

Jurěcková, J. and Sen, P.K. (1981). Sequential procedures based on M-estimators with discontinuous score functions. *J. Statist. Plan. Infer.* **5**, 253-66.

Jurěcková, J. and Sen, P.K. (1996). *Robust Statistical Procedures Asymptotics and Interrelations.* John Wiley & Sons, US.

Khan, S., and Saleh, A.K.Md.E. (2001). On the comparison of pre-test and shrinkage estimators for the univariate normal mean. *Stat. papers.* **42**, 451-473.

Khan, S., Hoque, Z., and Saleh, A.K.Md.E. (2002). Estimation of the slope parameter for linear regression model with uncertain prior information. *J. Stat. Res.* **36**, 55-74.

Saleh, A.K.Md.E. and Sen, P.K. (1982). Non-parametric tests for location after a preliminary test on regression. *Communications in Statistics: Theory and Methods.* **11**, 639-651.

Saleh, A.K.Md.E. (2006). *Theory of Preliminary test and Stein-type estimation with applications.* John Wiley & Sons, New Jersey.

Sen. P.K. (1982). On M-tests in linear models. *Biometrika.* **69**, 245-248.

Yunus, R.M. and Khan, S. (2007). Increasing power of the test through pre-test - a robust method. USQ Working Paper Series SC-MC-0713.

Wilcox, R.R. (2005). *Introduction to Robust Estimation and Hypothesis Testing.* Elsevier Inc, US.