Journal of Applied Probability and Statistics
2001x, Vol. xx, No. xx, pp. xx-xx

SCORE TEST IN ROBUST M-PROCEDURE

Shahjahan Khan[1]
School of Agricultural, Computational and Environmental Sciences
International Centre for Applied Climate Science
University of Southern Queensland, Toowoomba, Australia
Email: shahjahan.khan@usq.edu.au

Rossita M Yunus
Institute of Mathematical Sciences, Faculty of Sciences,
University of Malaya, Kuala Lumpur, Malaysia.
Email: rossita@um.edu.my

SUMMARY

A score type test based on the M-estimation method for a linear regression model is more reliable than the parametric based-test under mild departures from model assumptions, or when dataset has outliers. An R-function is developed for the score M-test, and applied to two real datasets to illustrate the procedure. The asymptotic power function of the M-test under a sequence of (contiguous) local alternatives is derived. Through computation of power function from simulated data, the M-test is compared with its alternatives, the Student's t and Wilcoxon's rank tests. Graphical illustration of the asymptotic power of the M-test is provided for randomly generated data from the normal, Laplace, Cauchy, and logistic distributions.

Keywords and phrases: Robust inference, M-test, Student's t test, rank test, asymptotic power, and contiguity.

2010 Mathematics Subject Classification: Primary 62F03 & 62F35, and Secondary 62G10

1 Introduction

Robust statistical methods are essential to avoid any misleading or devastating impact on the inference due to the presence of any outliers, and/or violation of model assumptions. This approach is crucial when the traditional assumptions on the parametric inference are not satisfied or there are outliers in the sample dataset. The validity of any statistical inference depends on the appropriateness of the method applied. The commonly used Student's

[1]The work was initiated when the first author was visiting University of Malaya.

t test (introduced by Gosset, (1908)) is heavily dependent on the assumption of normal populations, and as such it is not valid for the data obtained from any other distributions. Moreover, the presence of outliers in the data makes the Student's t test inappropriate.

Robust estimation methods are classified into three broad categories; M, L, and R-estimation (Huber, 1964). The M-estimation methods can be regarded as a generalization of maximum-likelihood estimation. The L-estimation methods are linear combinations of the ordered statistics, and the R-estimation methods are based on ranks of the observed data. Statistical tests were developed from/for the three categories of robust methods in the literature. For example, a signed rank test for a one sample location problem, a rank sum test for a two sample location problem (Wilcoxon, 1945), a rank test for linear models (Hájek, 1962, Saleh and Sen, 1983) are among the popular tests that are based on ranks of the observed data. Using the M-estimation method, some robust tests were proposed in the literature. For example, Schrader and Hettmansperger (1980) proposed a test based on the likelihood ratio criterion, Fung et al. (1985) proposed a test that kept the form of the Student's t-test but used the score function in the M-estimation method to make their proposed test robust. Sen (1982) introduced a score M-test for linear models; and Yunus and Khan (2010, 2011a, 2011b) used the score M-test to investigate the effect of the pre-testing on the slope parameter on the final testing of the intercept parameter of the linear regression models.

The nonparametric tests use the ranks of the observed data to formulate suitable test based on the rank sum statistics. In the process of ranking the observed data, valuable information, (details or magnitude) are lost, and is likely to impact on the quality of the test. For this reason the power of the nonparametric tests, in general, are lower than the equivalent parametric tests if the underlying distribution of the population is normal.

Any inference based on the M-procedure uses the original observed data values but treats the outliers to eliminate or minimize their impact on the inference using appropriate re-allocation of weights. As such, the M-procedure is less dependent on the assumptions of the population distribution. In the above sense, the M-test is robust. Although the exact distribution of the the M-test statistic is unavailable, its asymptotic distribution is used to workout the power function of the test. Since the M-statistic asymptotically follows a normal distribution its critical values are available from the normal table.

In the literature, Markatou and Hettmansperger (1990) generalized the Sen's score M-test to a bounded influence procedure. Heritier and Ronchetti (1994) and Silvapulle and Silvapulle (1995) are among others who studied along the line of the generalized score M-test. Some robust tests have robustness criteria of the class of the generalized M-estimators (GM) which down weight high leverage points using the Mallows-type weights (see works by Markatou and He (1994), Sinha and Wiens (2002), and Gagliardini et al. (2005)). The asymptotic distributions of the GM-test statistics however are somehow complicated (cf. Muller, 1998). As a result, we believe that it will be difficult to derive the asymptotic power function under a sequence of local alternative hypotheses for these test statistics. The concept of contiguity probability measures is used to derive the asymptotic distributions under the alternative hypotheses.

The derived asymptotic power functions of the score M-test are computed and graphically presented in this paper. Our paper provides the graphical analysis of the power function which was not the focus of many articles published earlier in this area (eg. Sen (1982) and Heritier and Ronchetti (1994)). The mathematical formula of power function is not reported in many articles in the area of robust statistical tests. Although the form of the power function of the score M-test is given in Sen (1982) and Jurečková and Sen (1996), the illustration of the power function through computation is unavailable. In our paper, the M-test for the one location and difference between two locations are implemented for real life data using newly defined R-functions. The associated t and rank tests are also accompanied and compared with the M-tests. The power function of the M-test is presented graphically, and it is compared to that of the Student's t and Wilcoxon's rank tests.

Since the M-test is not included in any popular statistical package, we provide the R-code and appropriate R-function to run the M-test for any given dataset for both one and two-sample cases. The R-function produces the observed value of the M-test statistic and the associated p-value. These values can be easily compared to the result of the relevant Student's t or relevant nonparametric test.

The two-sided M-test and its properties are given for one population and two populations in Section 2. Section 3 discusses the properties of the one-sided M-tests. Section 4 covers discussions on application of M-tests on two independent datasets. Section 5 illustrates the graphical comparisons of the power of the Student's t, Wilcoxon's rank and M-tests. The final Section provides discussions and concluding remarks. The R-codes and functions for the M-test are included in the Appendix.

2 The M-test

A linear regression model of n observable random variables, $Y_i,\ i = 1, \ldots, n$ is given by

$$Y_i = \theta + \beta x_i + e_i, \tag{2.1}$$

where the x_i's are known real constants of the explanatory variable with error term e_i, and θ and β are the unknown intercept and slope parameters respectively.

Let $Y_1, Y_2, \ldots, Y_n$ be identically and independently distributed random variables from a continuous distribution function $F(y) = F_i(y_i - \theta - \beta x_i),\ y_i, \theta, \beta \in \Re$. Also, assume that $F(y_i - \theta - \beta x_i)$ is a symmetric (about zero) distribution function.

M-estimators for θ and β are defined as the roots of the system of equations:

$$\sum_{i=1}^{n} \psi\left(\frac{Y_i - \theta - \beta x_i}{S_n}\right) = 0, \tag{2.2}$$

$$\sum_{i=1}^{n} x_i \psi\left(\frac{Y_i - \theta - \beta x_i}{S_n}\right) = 0, \tag{2.3}$$

where ψ is known as the score function in the M-estimation methodology. Here, S_n is an appropriate scale statistic for some functional $S = S(F) > 0$ and S_n is chosen to be the median of the absolute deviations of the sample from its median.

The choice of a suitable robust ψ-function justifies the test statistic. Several ψ-functions are available in the literature, among them the popular ones are the Huber's, Hampel's, and Tukey's ψ-functions.

- The Huber's score function is defined as $\psi_{Huber}(x) = x$ for $|x| \leq c$, c sign(x) for $|x| > c$, where x is any real number and c is known as the tuning constant because it can be chosen to fine tune the estimator. The value of the tuning constant is chosen as 1.345 since this value produces a 95% efficiency relative to the mean sample (Holland and Welsch, 1977).

- The Hampel's score function is written as $\psi_{Hampel}(x) = x,$ for $0 < |x| \leq a,$ a sign(x), for $a \leq |x| \leq b,$ $a(r - |x|)\text{sign}(x)/(r - b),$ for $b \leq |x| \leq r,$ 0 for $r \leq |x|$, where x is any real number and a, b and r are the tuning constants. The default values for these tuning constants used in R are $a = 2$, $b = 4$ and $r = 8$.

- The Tukey's score function is expressed as $\psi_{Tukey}(x) = x(1 - (x/k)^2)^2$ for $|x| \leq k$, 0 for $x > k$, where x is any real number and k is a tuning constant. The value of the tuning constant is chosen as 4.685 since this value produces a 95% efficiency relative to the mean sample (Holland and Welsch, 1977).

In this paper, we consider three special cases of hypotheses testing in the linear regression model, (i) testing location of a population distribution, (ii) testing the equality of the locations of two population distributions and (iii) testing on the slope coefficient of a regression model.

2.1 The M-test for one location

Let $\beta = 0$ in the equation (2.1), so θ is the location of the distribution of Y. Assume that the distribution $F(y - \theta)$ is continuous. We wish to test the location of the distribution to be a specified value, that is, $H_0 : \theta = \theta_0$ against $H_0 : \theta \neq \theta_0$.

An appropriate M-test, to test $H_0 : \theta = \theta_0$ against $H_A : \theta \neq \theta_0$, is based on the following test statistic

$$M_{1n} = M_{1n}(\theta_0) = \sum_{i=1}^{n} \psi\left(\frac{Y_i - \theta_0}{S_{1n}}\right) \tag{2.4}$$

with scale statistic S_{1n}. At the α-level of significance, the H_0 is rejected if the observed value of the test statistic satisfies $|M_{1n}| > \ell_{1n,\alpha/2}$, where $\ell_{1n,\alpha/2}$ is the upper $\alpha/2$-percentile of the distribution of M_{1n}.

According to Sen (1996), an M-estimator is consistent if the ψ-function is bounded and skew symmetric, and the true distribution of the population is symmetric. If $F(y - \theta)$ is symmetric about zero, then

$$\int_{-\infty}^{\infty} \psi \left(\frac{Y_i - \theta_0}{S_{1n}} \right) dF(Y_i - \theta_0) = 0.$$

For large samples, under $H_0 : \theta = \theta_0$,

$$n^{-\frac{1}{2}} M_{1n}(\theta_0)/S^{\star}_{1n} \to N(0, 1), \tag{2.5}$$

where $S^{*^2}_{1n} = n^{-1} \sum_{i=1}^{n} \psi^2 \left(\frac{Y_i - \tilde{\theta}}{S_{1n}} \right)$ in which $\tilde{\theta}$ is the studentized M-estimator of θ based on $Y_1, Y_2, \ldots, Y_n$, and it is expressed as

$$\tilde{\theta} = \frac{1}{2} \sup \left\{ a : \sum_{i=1}^{n} \psi((Y_i - a)/S_{1n}) > 0 \right\} + \frac{1}{2} \inf \left\{ a : \sum_{i=1}^{n} \psi((Y_i - a)/S_{1n}) < 0 \right\},$$

and $S^{*^2}_{1n} \to \sigma_1^2$ as $n \to \infty$, (cf. Jurečková and Sen, 1996, p. 409) and

$$\sigma_1^2 = \int_{\Re} \psi^2 \left(\frac{Y_i - \theta}{S_1} \right) dF(Y_i - \theta), \quad (0 < \sigma_1 < \infty)$$

is the second moment of $\psi(\cdot)$. If $\psi(x) = x$ (i.e. the maximum likelihood ψ-function) and $F \sim N(0, \sigma^2)$, then $S_1 = \sigma$ and $\sigma_1^2 = 1$.

2.1.1 Properties of M_{1n}

Let α be the nominal significance level for the above test. Then the critical value $\ell_{1n,\alpha/2}$ is such that

$$P(|M_{1n}| > \ell_{1n,\alpha/2}|H_0) = P(M_{1n} > \ell_{1n,\alpha/2}, M_{1n} < -\ell_{1n,\alpha/2}|H_0) = \alpha. \tag{2.6}$$

Let $\phi_{1n} = I(|M_{1n}| > \ell_{1n,\alpha/2})$ be the test function designated to test $H_0 : \theta = \theta_0$ against $H_A : \theta \neq \theta_0$, where $I(A)$ is the indicator function of the set A which assumes values 0 or 1. Also, let $\Phi(x)$ be the standard normal distribution function of the random variable X and $\Phi(\tau_{\alpha/2}) = 1 - \alpha/2$, $0 < \alpha < 1$. From equations (2.5) and (2.6), as $n \to \infty$,

$$n^{-\frac{1}{2}} \ell_{1n,\alpha/2}/S^{*}_{1n} \to \tau_{1\alpha/2}. \tag{2.7}$$

Now let $\alpha_{1n} = E(\phi_{1n}|\theta = \theta_0)$ be the size of ϕ_{1n}. Then,

$$\alpha_{1n} = P(|M_{1n}| > \ell_{1n,\alpha/2}|H_0 : \theta = \theta_0) = \alpha$$

using equation (2.6). The power function of the test function ϕ_{1n}, is defined as

$$\begin{aligned} \Pi_{1n}(\theta) &= E(\phi_{1n}|\theta) = P(|M_{1n}| > \ell_{1n,\alpha/2}| \text{ any } \theta) \\ &= P\left[M_{1n} > \ell_{1n,\alpha/2}| \text{ any } \theta\right] + P\left[M_{1n} < -\ell_{1n,\alpha/2}| \text{ any } \theta\right]. \end{aligned} \tag{2.8}$$

Note that the size of the test α_{1n} is a special case of the power function of the test when the null hypothesis is true, i.e. $\alpha_{1n} = \Pi_{1n}(\theta = \theta_0)$.

From the equation (5.5.29) of Jurečková and Sen (1996, p. 221), under H_0, as n grows large,

$$\sup\left\{n^{-\frac{1}{2}}|M_{1n}(\theta_0 + a) - M_{1n}(\theta_0) + n\gamma_1 a| : |a| \le n^{-\frac{1}{2}}K\right\} \to 0, \tag{2.9}$$

where K is a positive constant, and

$$\gamma_1 = \frac{1}{S_1}\int_{\Re} \psi'\left(\frac{Y_i - \theta}{S_1}\right) dF(Y_i - \theta)$$

in which ψ' is the derivative of ψ-function.

Further, consider a sequence of local alternative hypotheses $\{H_n\}$, where

$$H_n : \theta = \theta_0 + n^{-\frac{1}{2}}\lambda, \quad \lambda > 0.$$

Now utilizing the contiguity of probability measures (see Hájek et al., 1999, Ch. 7) under $\{H_n\}$ to those under H_0, equation (2.9) implies that $n^{-\frac{1}{2}}M_{1n}(\theta_0)$ under $\{H_n\}$ is asymptotically equivalent to $n^{-\frac{1}{2}}M_{1n}(\theta_0 + n^{-\frac{1}{2}}\lambda) + \lambda\gamma_1$. However, the asymptotic distribution of $n^{-\frac{1}{2}}M_{1n}(\theta_0)$ under $\{H_n\}$ is the same as the distribution of $n^{-\frac{1}{2}}M_{1n}(\theta_0 - n^{-\frac{1}{2}}\lambda) = n^{-\frac{1}{2}}M_{1n}(\theta_0) + \lambda\gamma_1$ under H_0, by the fact that the distribution of $M_{1n}(a)$ under $\theta = a$ is the same as that of $n^{-\frac{1}{2}}M_{1n}(\theta - a)$ under $\theta = 0$ (cf. Saleh, 2006, p. 332). Therefore, for a large sample, under $\{H_n\}$ the distribution of

$$n^{-\frac{1}{2}}M_{1n} \to N(\lambda\gamma_1, \sigma_1^2). \tag{2.10}$$

Thus, under $\{H_n\}$, the asymptotic power function of the one-sample M-test is given by

$$\Pi_1(\lambda) = \lim_{n\to\infty} \Pi_{1n} = 1 - \Phi(\tau_{1\alpha/2} - \lambda\gamma_1/\sigma_1) + \Phi(-\tau_{1\alpha/2} - \lambda\gamma_1/\sigma_1) \tag{2.11}$$

using equations (2.8) and (2.10). Obviously, for any large sample size, the asymptotic size of the test for one-sample M-test is given by

$$\alpha_1 = \Pi_1(\lambda = 0) = 1 - \Phi(\tau_{1\alpha/2}) + \Phi(-\tau_{1\alpha/2}) = \alpha.$$

2.2 The M-test for difference of two locations

Let two independent random samples, $U_1, U_2, \ldots, U_{n_1}$, and $V_1, V_2, \ldots, V_{n_2}$, be drawn from the populations of U and V such that,

$$P(U_i \le t) = P(V_j \le t + \beta) = F(t), \tag{2.12}$$

where $F(t)$ is the cumulative distribution function (cdf) of a continuous distribution. Thus, the two locations differ by a constant β, that is, the location of the distribution of V is shifted by β from the distribution of U.

We want to test H_0^*: distribution of U and V are identical against H_A^*: V has different location than U, and this is equivalent to test $H_0^* : \beta = 0$ against $H_A^* : \beta \neq 0$.

Let the two random samples from U and V be merged to form a combined random sample $Y_1, Y_2, \ldots, Y_n$, such that $Y_1 = U_1,\ Y_2 = U_2, \ldots, Y_{n_1} = U_{n_1}, Y_{n_1+1} = V_1, \ldots, Y_n = V_{n_2}$, where $n = n_1 + n_2$. Then the predictor variable in the equation (2.1) $x_k = 0$, for $k = 1, 2, \ldots, n_1$, and $x_k = 1$, for $k = n_1 + 1, \ldots, n$.

Consider a M statistic,

$$M_{2n}^*(\tilde{\theta}, 0) = \sum_{j=1}^{n_2} \psi\left(\frac{V_j - \tilde{\theta}}{S_{2n}}\right) = \sum_{k=1}^{n} x_k \psi\left(\frac{Y_k - \tilde{\theta}}{S_{2n}}\right), \tag{2.13}$$

where ψ is the score function and S_{2n} is an appropriate scale statistic for some functional $S_2 = S_2(F) > 0$. The median of the absolute deviations of the sample Y from its median is used as an estimate of S_{2n}. Note that $\tilde{\theta}$ is the constrained M-estimator of θ when $\beta = 0$, that is, $\tilde{\theta}$ is the solution of $\sum_{k=1}^{n} \psi\left(Y_k - a\right) = 0$ and conveniently be expressed as

$$\tilde{\theta} = \frac{1}{2}\sup\left\{a : \sum_{k=1}^{n} \psi\left(\frac{Y_k - a}{S_{2n}}\right) > 0\right\} + \frac{1}{2}\inf\left\{a : \sum_{k=1}^{n} \psi\left(\frac{Y_k - a}{S_{2n}}\right) < 0\right\}.$$

From Sen (1982) and Yunus and Khan (2011a), under $H_0^* : \beta = 0$,

$$M_{2n} = \frac{1}{S_{2n}^* \sqrt{n_1 n_2 / n}} \sum_{j=1}^{n_2} \psi\left(\frac{V_j - \tilde{\theta}}{S_{2n}}\right) \to N(0,1) \text{ as } n \to \infty, \tag{2.14}$$

where $S_{2n}^{*^2} = n^{-1}\left[\sum_{k=1}^{n} \psi^2\left(\frac{Y_k - \tilde{\theta}}{S_{2n}}\right)\right]$.

2.2.1 Properties of M_{2n}

Consider a local sequence of alternative hypotheses $\{K_n\}$, where

$$K_n : \beta = n^{-\frac{1}{2}}\eta, \quad \eta > 0. \tag{2.15}$$

Following similar steps as in the one-sample case, the asymptotic power function of the M-test for the two-sample problem under $\{K_n\}$ is given by

$$\Pi_2(\eta) = \lim_{n\to\infty} \Pi_{2n}(\beta) = 1 - \Phi(\tau_{2\alpha/2} - \eta\gamma_2\sqrt{n_1 n_2}/n\sigma_2) + \Phi(-\tau_{2\alpha/2} - \eta\gamma_2\sqrt{n_1 n_2}/n\sigma_2) \tag{2.16}$$

using the asymptotic results of Jurečková and Sen (2006), and Yunus and Khan (2011a), where

$$\gamma_2 = \frac{1}{S_2}\int_{\Re} \psi'\left(\frac{Y_k - \theta - \beta x_k}{S_2}\right) dF(Y_k - \theta - \beta x_k), \tag{2.17}$$

in which ψ' is the derivative of the ψ-function and

$$\sigma_2^2 = \int_{\Re} \psi^2\left(\frac{Y_k - \theta - \beta x_k}{S_2}\right) dF(Y_k - \theta - \beta x_k), \quad (0 < \sigma_2 < \infty) \tag{2.18}$$

is the second moment of $\psi(\cdot)$.

The asymptotic size of the test is given by

$$\alpha_2 = \Pi_2(0) = 1 - \Phi(\tau_{2\alpha/2}) + \Phi(-\tau_{2\alpha/2}) = \alpha \tag{2.19}$$

from equation (2.16).

2.3 The M-test for testing the slope coefficient

A convenient form of the M-test statistic for testing $H_0 : \beta = \beta_0$ against $H_A : \beta \neq \beta_0$ for the model in (2.1) is given by

$$M_n = M_n^\star(\tilde{\theta}_m, \beta_0) = \sum_{i=1}^{n} x_i \psi\left(\frac{Y_i - \tilde{\theta}_m - \beta_0 x_i}{S_n}\right) \text{ say,} \tag{2.20}$$

where $\tilde{\theta}_m$ is the constrained M-estimator of θ when $\beta = \beta_0$, that is, $\tilde{\theta}_m$ is the solution of $M_n^\star(a, \beta_0) = 0$ and it may be conveniently be expressed as

$$\tilde{\theta} = [\sup\{a : M_n^\dagger(a, \beta_0) > 0\} + \inf\{a : M_n^\dagger(a, \beta_0) < 0\}] \div 2, \tag{2.21}$$

where $M_n^\dagger(a, b) = \sum_{i=1}^{n} \psi\left(\frac{Y_i - a - bx_i}{S_n}\right)$, a and b are any real numbers. Then H_0 is rejected if $|M_n| > \ell_{n,\alpha/2}$ at the α level of significance, where $\ell_{n,\alpha/2}$ is the upper $\alpha/2$-percentile of the distribution of M_n.

It follows from the equation (2.6) of Yunus and Khan (2011a) that under H_0,

$$n^{-\frac{1}{2}} M_n \xrightarrow{d} N(0, \sigma_0^2 C^{\star 2}) \text{ as } n \to \infty, \tag{2.22}$$

where $C^\star = \lim_{n\to\infty} \sum_{i=1}^{n} x_i^2 - n\bar{x}_n^2$, $\bar{x}_n = n^{-1}\sum_{i=1}^{n} x_i$, and
$\sigma_0^2 = \int_{\Re} \psi^2 \left(\frac{Y_i - \theta - \beta x_i}{S}\right) dF(Y_i - \theta - \beta x_i)$ $(0 < \sigma_0 < \infty)$
is the second moment of $\psi(\cdot)$. Let

$$S_n^{\star 2} = n^{-1} \sum_{i=1}^{n} \psi^2 \left(\frac{Y_i - \tilde{\theta}_m - \beta_0 x_i}{S_n}\right), \tag{2.23}$$

(cf. Jurečková and Sen, 1996, p. 409) and $S_n^{\star 2} \to \sigma_0^2$ as $n \to \infty$.

2.4 Properties of M_n

Now consider a sequence of local alternative hypotheses $\{Q_n\}$, where

$$Q_n : \beta = \beta_0 + n^{-\frac{1}{2}}\nu, \nu > 0. \tag{2.24}$$

Using equations (2.22), and (5.5.29) of Jurečková and Sen (1996), and the contiguity probability measures, under $\{Q_n\}$, the distribution of $n^{-\frac{1}{2}} M_n \xrightarrow{d} N(\nu\gamma C^{\star 2}, \sigma_0^2 C^{\star 2})$ (cf. Yunus and Khan, 2011a).

Following similar steps as in the one-sample case, the asymptotic power function of the M-test for testing the slope coefficient of the regression model under $\{Q_n\}$, is given by

$$\Pi_M(\nu) = 1 - \Phi(\tau_{\alpha/2} - \nu\gamma C^\star \sigma_0^{-1}) + \Phi(-\tau_{\alpha/2} - \nu\gamma C^\star \sigma_0^{-1}), \tag{2.25}$$

where

$$\gamma = \frac{1}{S}\int_{\Re} \psi'\left(\frac{Y_i - \theta - \beta x_i}{S}\right) dF(Y_i - \theta - \beta x_i)$$

and ψ' is the derivative of ψ-function. Here $\tau_{\alpha/2}$ is the critical value of the standard normal distribution at the $\alpha/2$ level of significance.

3 The one-sided tests

The adoption of the above two-sided M-test to a one-sided test is straightforward. Suppose we test $H_0 : \beta = 0$ against $H_A : \beta > 0$, then we work with M_{1n} and the corresponding critical value $\ell_{1n,\alpha}$ in (2.7), and obtain $P(M_{1n} > \ell_{1n,\alpha}|H_0) = \alpha$. It follows that the power function for a one-sided M-test for testing the location of one population is given by

$$\Pi_{1n}(\theta) = P\left[M_{1n} > \ell_{n,\alpha}|\text{ any } \theta\right]. \tag{3.1}$$

As $n \to \infty$, we find that the asymptotic power of a one-sided test for testing about the location of population is given by

$$\Pi_1(\lambda) = \lim_{n\to\infty} \Pi_{1n}(\mu) = 1 - \Phi(\tau_{1\alpha} - \lambda\gamma_1\sigma_1^{-1}). \tag{3.2}$$

In the same manner, the asymptotic power of a one-sided test for testing the equality of location parameter of two populations is given by

$$\Pi_2(\eta) = 1 - \Phi(\tau_{2\alpha} - \eta\gamma_2\sqrt{n_1 n_2}/n\sigma_2) \tag{3.3}$$

and that for testing on the slope coefficient is given by

$$\Pi_M(\nu) = 1 - \Phi(\tau_\alpha - \nu\gamma C^\star \sigma_0^{-1}). \tag{3.4}$$

4 Applications on data

In this section, the R-codes to compute the value of the M-test statistic, its p value, confidence interval for θ and asymptotic power of the test are discussed. In the Appendix the R-function for the M-test is included. Users can choose to run a one-sided or two-sided test. The R-codes that produce the M-test statistic, p-value, confidence interval, σ_1 and γ_1 for one-sample test are given in Listing 1, while the R-codes that give the asymptotic power of the two-sample M-test are given in Listing 2. Examples of how to use the proposed

R-functions for one-sample M-test are given in Listing 3. For the two-sample M-test, the R-codes that produce the M-test statistic and its p-value, the asymptotic power and examples of using the M-test on data are given in Listings 4, 5, and 6. The formula and R-codes in this paper are based on the Tukey's ψ-function. The M-test for both one and two locations as well as its applications on two different datasets are included here.

4.1 One-sample M-test: Birth rate of 56 states in United States in 2010

For the illustration of the M-test of one location we consider the birth rate dataset obtained from National Vital Statistics Report (Table 12. Birth rates, by age of mother: United States, each state and territory, 2010). Birth rate of 56 states in United States were measured for year 2010. The mean and median of birthrate is 13.38 and 12.6 respectively. It is observed from the normal Q-Q plot and the histogram given in Figures 1(a) and (b) that the distribution of the data is not normal. In facts, these figures reveal some outliers.

The observed value of the test statistic for the t-, R- and M-tests (Student's t, Wilcoxon's rank and M-tests) along with the p-values are calculated for testing $H_0 : \mu = 13.5$ against $H_A : \mu \neq 13.5$ at the 5% significance level and are given in Table 1. We find that the t-test could not reject H_0 at the 5% significance level as the p-value is 0.7143. However, the R- and M-tests reject the null hypothesis as the p-values are 0.0557 and 0.0338, respectively (see Table 1). For this dataset, one may have a different null hypothesis, that is, to test μ at a particular value, say μ_0, as μ_0 can take any real number in this two-sided testing. We obtain p-value for each testing on the $H_0 : \mu = \mu_0$, and then we plot p-value against μ_0 in Figure 1(c). We observed that the M-test is comparable in performance to the R-test, but not to the t-test. Figure 1(d) shows the asymptotic power curves of the M- and t-tests for the birth rate dataset. Obviously, asymptotic power of the M-test is higher than that of the t test. Existing R-codes wilcox.test and t.test were used to find the statistics and p-values for the R- and t-tests, respectively, while coding for the M-test is given in the appendix.

Table 1: Test results for the birth rate data

t-test		R-test		M-test	
T-statistic	p-value	R-statistic	p-value	M-statistic	p-value
-0.368	0.7143	499	0.0557	-14.069	0.0338

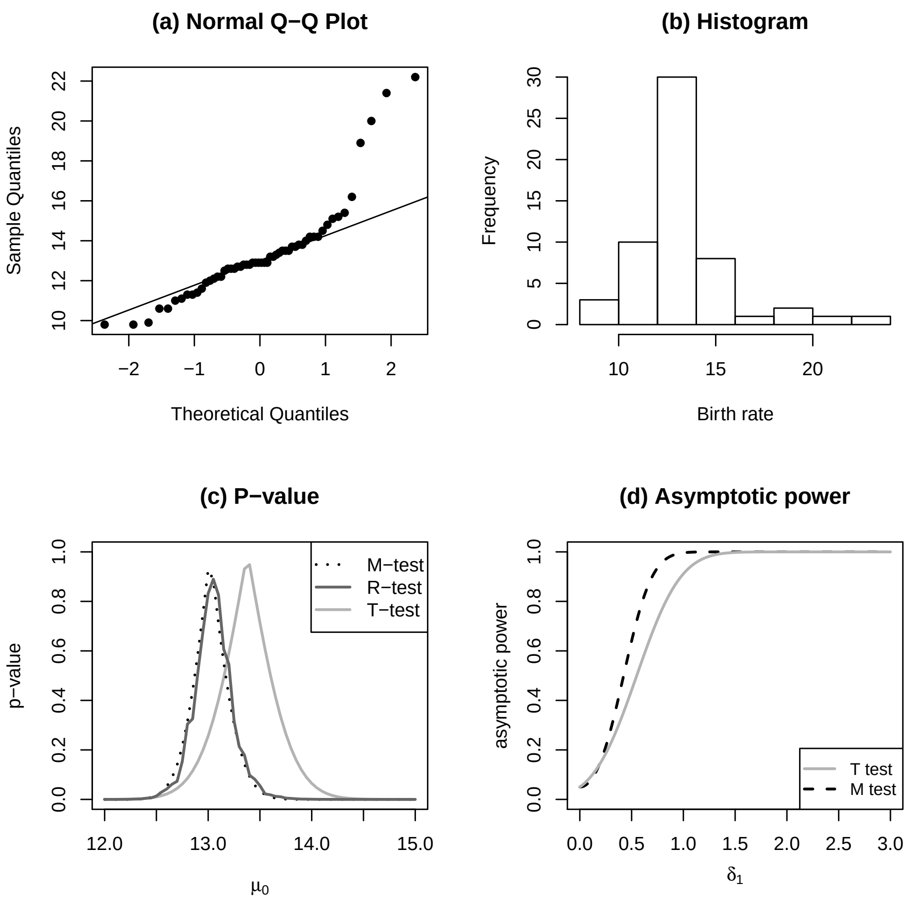

Figure 1: Graphs of Q-Q plot and asymptotic power curves for birth rate dataset, where $\delta_1 = |\mu - \mu_0|$.

Table 2: Test results for the iodine versus LOCM data

t-test		R-test		M-test	
T-statistic	p-value	R-statistic	p-value	M-statistic	p-value
0.56	0.5783	0.80	0.4247	0.77	0.4425

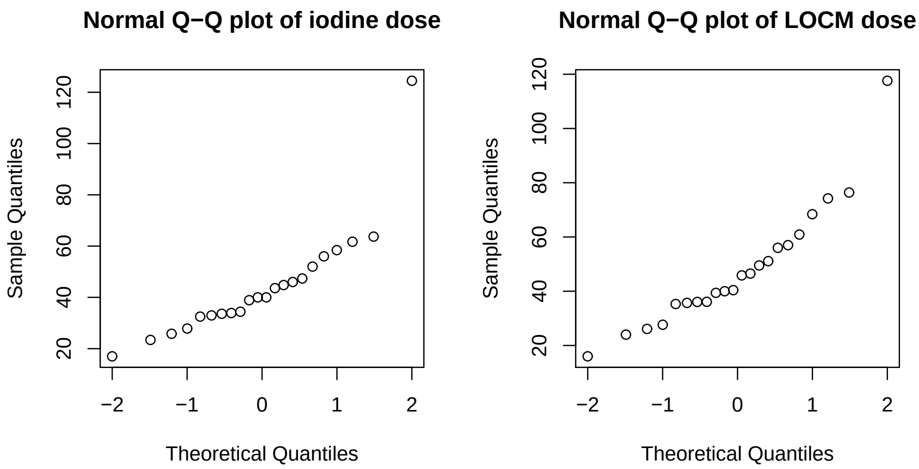

Figure 2: Graphs of Q-Q plot and histogram for iodine dose and LOCM dose.

4.2 Two-sample M-test: Iodine versus LOCM

A nephrotoxicity of iso-osmolar iodixanol is compared with a nonionic low-osmolar contrast media (LOCM) to find out which of them is more effective in reducing the risk of contrast media-induced nephropathy. In the study by Heinrich et al. (2009), serum creatinine levels are assessed before and after an intervascular application of iodixanol and LOCM.

The average of iodine and LOCM dose (mg/dL) are taken from 22 studies. We consider to test H_0^*: the distributions of iodine dose and LOCM dose are identical against H_A^*: the location of the distribution of iodine dose is different from the location of the distribution of LOCM dose. It is observed that there is one outlier in each normal Q-Q plot for the iodine dose and LOCM dose (see Figure 2). In the testing, we find H_0^* is not rejected at the 5% significance level using the t, rank and M-tests, respectively with p-values 0.5783, 0.4247 and 0.4425 (see Table 2).

5 Power Comparison

In this Section, simulated data sets is used to obtain the asymptotic power of the M-test for situations in which samples were drawn from several symmetrical distributions. The asymptotic power of the M-test is compared to those of the commonly used t- and R-tests in the simulation.

5.1 Test on the location of a population

Consider $X_i = \mu + e_i$, $i = 1, 2, \ldots, n$ where μ is the location parameter and X_i is a random response with error e_i. For the simulation, wet set $\mu = 2$, $\alpha = 0.05$, and $n = 100$.

Four symmetric distributions, namely the (i) normal, (ii) Laplace, (iii) Cauchy, and (iv) logistic, of error terms are considered to compare the asymptotic power of the tests. For the normal case, e_i is generated from a normal distribution with mean 0 and variance 1. For the Laplace and Cauchy cases, e_i is generated respectively from a Laplace and Cauchy distribution with location 0 and scale 1, while for the logistic case, e_i is from logistic distribution with location 0 and scale $1/\sqrt{3}$.

Asymptotic power of the M-test is computed using the function given in the equation (2.11) for the two-sided test. The estimate of γ_1 in the equation (2.11) is taken as $\hat{\gamma}_1 = \frac{1}{n\ (MAD/0.6745)} \sum_{i=1}^{n} \psi'_{huber}\left(\frac{X_i - \tilde{\mu}}{MAD/0.6745}\right)$, where MAD is the median absolute deviation of the sample of X. The σ_1 in the equation (2.11) is estimated by S_{1n}^* using $\psi = \psi_{huber}$.

The simulation is run 10,000 times to get 10,000 simulated sets of values of error terms. Using $X_i = 2 + e_i$, $i = 1, 2, \ldots, n$, we obtain 10,000 simulated datasets of size $n = 100$. Then, these datasets are used to compute S_{1n}^* and $\hat{\gamma}_1$. The average of asymptotic power of the test for the 10,000 simulated datasets is computed at a particular value of $\delta_1 = |\mu - \mu_0|$. After 10,000 repetitions, the value of δ_1 was increased and the process repeated. The curves of the asymptotic power of the tests for increasing values of δ_1 are plotted in Figure 3.

It is depicted in Figure 3(a) that asymptotic power of the M-test is as much as that of the t-test, and power of both tests are slightly higher than that of the R-test when data is generated from normal distribution. However, the asymptotic power of the R- and M-tests is larger than that of the t-test when sample data is generated from the Laplace and Cauchy distributions ((b) and (c)). It is observed that M-test is comparable in terms of power to the R-test when the distribution of data is Cauchy (heavy tails) or Laplace (light tails). All the tests have similar power when sample data is generated from logistic distribution ((d)).

5.2 Test on the equality of location of two populations

Consider two independent random samples, $U_1, U_2, \ldots, U_{n_1}$ and $V_1, V_2, \ldots, V_{n_2}$, from the random variables U and V, where the two distributions are identical except for the difference in the location. Let β be the difference between the two locations of the two populations. In the simulation study, we set $\alpha = 0.05$ and $n_1 = n_2 = 100$, so $n = n_1 + n_2 = 200$.

Four distributions, namely the (i) normal, (ii) Laplace, (iii) Cauchy, and (iv) logistic, of U and V are considered to compute the asymptotic power function of the M-test. For the normal case, U is generated from a normal distribution with location/mean 2 and variance 1 and V is generated from a normal distribution with location/mean $2 + \beta$ and variance 1. For the Laplace, Cauchy and logistic cases, U is generated respectively from a Laplace, Cauchy or logistic distribution with a location parameter 2 and a scale parameter 1 and V is generated respectively from a Laplace, Cauchy, or logistic distribution with a location parameter $2 + \beta$ and a scale parameter 1. Obviously the two samples of U and V have

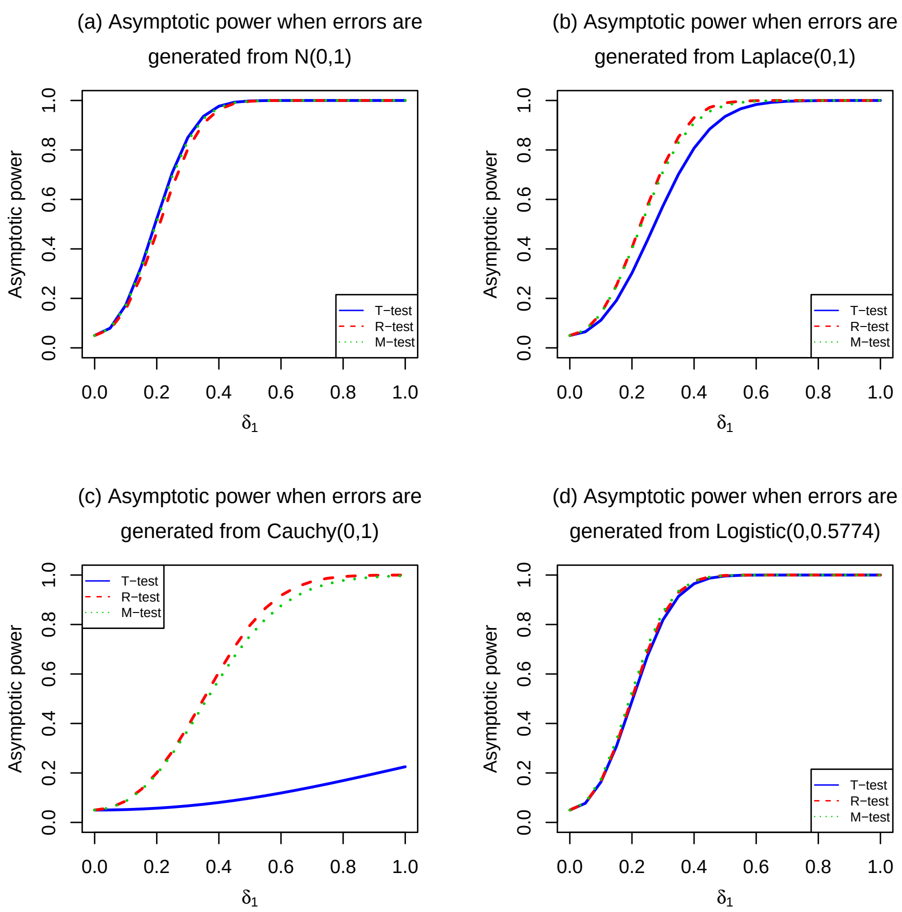

Figure 3: Graphs of the power function for increasing δ_1, where $\delta_1 = |\mu - \mu_0|$ at $\alpha = 0.05$.

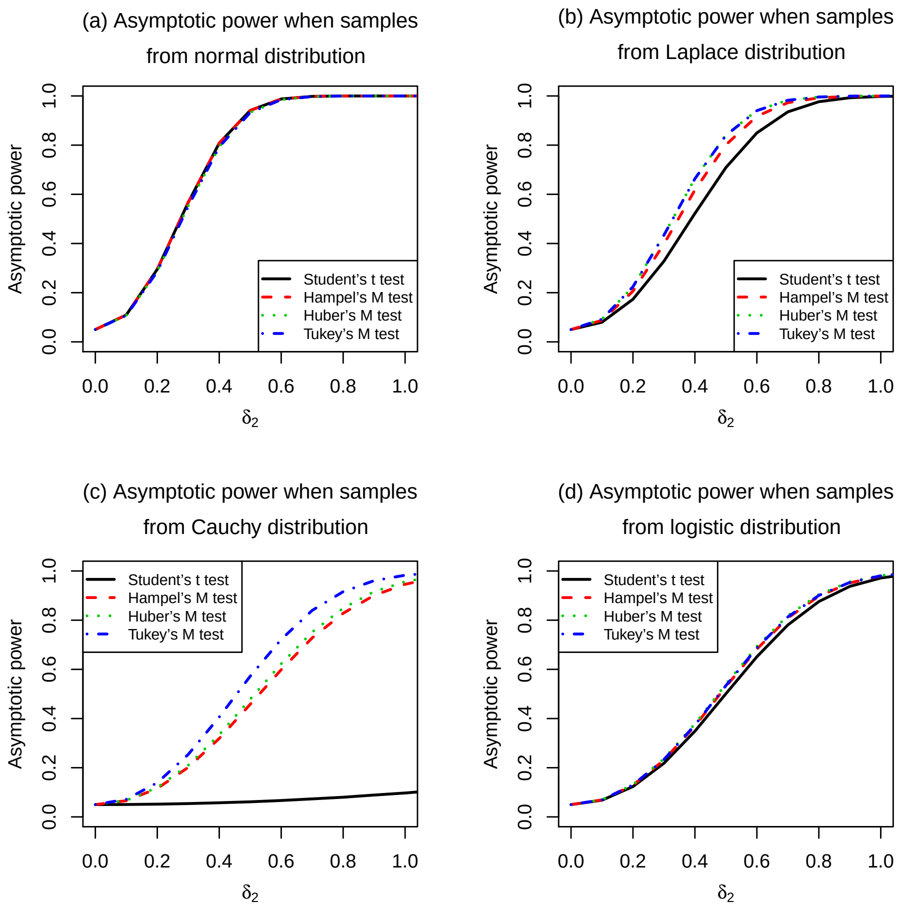

Figure 4: Graphs of the power function for increasing δ_2, where $\delta_2 = |\beta - \beta_0|$ at $\alpha = 0.05$.

identical distribution if $\beta = 0$.

Asymptotic power of the two-sided M-test is computed using the form of asymptotic power function given in equation (2.16). The estimate of γ_2 in the equation (3.3) is taken as $\frac{1}{n\, MAD/0.6745} \sum_{k=1}^{n} \psi' \left(\frac{Y_k - \hat{\theta}_m - \hat{\beta}_m c_k}{MAD/0.6745} \right)$ and σ_2 is estimated by $\sqrt{\frac{1}{n} \sum_{k=1}^{n} \psi^2 \left(\frac{Y_k - \hat{\theta}_m - \hat{\beta}_m c_k}{MAD/0.6745} \right)}$, where $\hat{\theta}_m$ and $\hat{\beta}_m$ are the M-estimates for parameters θ and β of the simple regression model in (2.10), and MAD is the median absolute deviation of the sample of Y. In the simulation, the Hampel's, Huber's, and Tukey's ψ-functions are considered to obtain the asymptotic power of the M-test.

The simulation is run 10,000 times to get 10,000 simulated sets of values of both samples datasets. Then, these datasets are used to compute S_{2n}^*, and $\hat{\gamma}_2$. After 10,000 repetitions, the value of the δ_2 was increased and the process repeated. The asymptotic power curves for increasing values of δ_2 were plotted in Figure 4.

From Figure 4, we find that the M-test based on the Hampel's, Huber's, and Tukey's ψ-functions are more robust against departures from the normal distribution assumption as their powers are larger than that of the Student's t-test. The power of M-test based on the Hampel's ψ-function are close to that of the Student's t-test when sampling is done from the normal distribution. The M-test based on the Huber's and Tukey's ψ-functions have larger power than that of the Hampel's when the samples are from the Laplace and Cauchy distributions.

6 Concluding remarks

The use of M-test removes any chance of misleading test outcome due to the violation of assumptions or existence of outliers. Furthermore, the asymptotic power of the M-test is at least as large as that of the Student's t or relevant nonparametric test when the assumptions are not met and even if there are no outliers. Clearly for the users it is advantageous to use the M-test to avoid any risk of using a test whose underlying assumptions may have been violated and hence the validity of the test outcomes becomes untenable.

In many cases the ordinary users of statistical tests do not bother to check the validity of the assumptions. For those users M-test is a better option as it provides much needed protection against the adverse consequences of the presence of outliers or departure from the assumptions on the population distribution.

6.1 Acknowledgements

The authors are thankful to the editor and referees for their valuable suggestions and comments. The work of the second author was partially funded by the University of Malaya Research Grant (UMRG), no. RG169/11AFR, when the study was conducted.

References

[1] Fung K.Y., Lee H. and Tajuddin, I. (1985). Some Robust Test Statistics for the Two-Sample Location Problem. *Journal of the Royal Statistical Society Series D (The Statistician),* **34***, 2, 175-182.*

[2] Gagliardini, P., Trojani, F. and Urga, G. (2005). Robust GMM tests for structural breaks. *Journal of Econometrics,* **129***, 139-182.*

[3] Gosset, W.S. (1908). The probable error of a mean. *Biometrika,* **6***,125.*

[4] Hájek, J. (1962). Asymptotically Most Powerful Rank-Order Tests. *Ann. Math. Statist.,* **33***, 3 , 1124-1147.*

[5] Hájek, J., Šidák, Z. and Sen, P.K. (1999). *Theory of Rank Tests.* Academia Press, New York.

[6] Heritier, S. and Ronchetti, E. (1994), Robust bounded-influence tests in general parametric models. *J. Amer. Statist. Assoc.,* **89***, 897-904.*

[7] Holland, P.W. and Welsch, R.E. (1977). Robust regression using iteratively reweighted least-squares. *Communications in Statistics - Theory and Methods, A,* **6***, 813-838.*

[8] Huber, P.J. (1964). Robust estimator of a location parameter. *Ann. Math. Statist.,* **35***, 1, 73-101.*

[9] Jurečková, J. and Sen, P.K. (1996). *Robust Statistical Procedures Asymptotics and Interrelations.* Wiley, New York.

[10] Markatou, M. and Hettmansperger, T. P. (1990), Robust bounded-influence tests in linear models. *J. Amer. Statist. Assoc.,* **85***, 187-190.*

[11] Markatou M. He. X (1994) Bounded influence and high breakdown point testing procedures in linear models. *J. Amer. Statist. Assoc.,* **89***, 42, 543-549.*

[12] Muller, C. (1998). Optimum robust testing in linear models. *Ann. Statist.,* **26***, 1126-1146.*

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

[14] Saleh, A.K.Md.E. and Sen, P.K. (1983). Nonparametric tests of location after a preliminary test on regression in the multivariate case. *Communications in Statistics: Theory and Methods,* **12***, 1855-1872.*

[15] Schrader, R.M. and Hettmansperger, T.P. (1980). Robust analysis of variance based upon a likelihood ratio criterion. *Biometrika,* **67***, 1, 93-101.*

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

[17] Sen, P.K. (1996). Estimating Functions: Nonparametrics and Robustness. Ishwar V. Basawa, V. P. Godambe, and Robert L. Taylor, eds. Selected Proceedings of the Symposium on Estimating Functions: Held at the University of Georgia, Athens, GA, March 21–23, 1996 (Hayward, CA: Institute of Mathematical Statistics, 1997), 417-436.

[18] Silvapulle M.J and Silvapulle P. (1995). A Score Test Against One-Sided Alternatives. *J. Amer. Statist. Assoc.,* **90***, 429, 342 - 349.*

[19] Sinha S. and Wiens D.P. (2002) Minimax Weights for Generalised M-Estimation in Biased Regression Models. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique,* **30***, 3, 401-414.*

[20] Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin,* **1***, 6, 80-83.*

[21] Yunus, R.M. and Khan, S. (2010). Increasing power of robust test through pre-testing in multiple regression model. *Pakistan Journal of Statistics, Special Volume on 25th Anniversary, (Edited by S.E. Ahmed)* **26***, 151-170.*

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

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

R-Codes

```
(i) M-test for testing the location of a population

m.test1<-function(X, alternative = c("two.sided", "less", "greater"),
            mu.not, sig.level){
    n<-length(X)
    library(MASS)
    mad.X<-mad(X)
    fit<-rlm(X~1)
    r1<-(X-mu.not)/mad.X
    r2<-(X-fit$coef)/mad.X
    Mstat<-sum(psi.huber(r1,deriv=0)*r1)
    sigma <-sqrt((1/n)*sum((psi.huber(r2,deriv=0)*r2)^2))
    gamma <-(1/n)*(sum(psi.huber(r2, deriv = 1)))/mad.X
    standardized.Mstat<-Mstat/(sigma*sqrt(n))
        if (alternative =="greater")
        {
        p.value<-1-pnorm(standardized.Mstat)
        }
        if  (alternative =="less")
        {
        p.value <-pnorm(standardized.Mstat)
        }
        if (alternative =="two.sided"){
        p.value <-if (standardized.Mstat>=0)
                    2*(1-pnorm(standardized.Mstat)) else
```

```
                  2*pnorm(standardized.Mstat)
        }
    interval <-c(fit$coef-(qnorm(1-sig.level/2))*sqrt(1/n)*mad.X,
                 fit$coef+(qnorm(1-sig.level/2))*sqrt(1/n)*mad.X)
    list(Mstat=Mstat, standardized.Mstat=standardized.Mstat,
    p.value=p.value,M.estimate = fit$coef, interval=interval,
    sigma=sigma,     gamma=gamma)
    }
}
```

(ii) The asymptotic power of the M-test for testing the location of one population

```
power.m.test1<-function(n, alternative=c("one.sided","two.sided"),delta,
               sigma, gamma, sig.level){
    lambda<-delta*sqrt(n)
    if (alternative =="one.sided"){
        power<-1-pnorm(qnorm(1-sig.level)- lambda*gamma/sigma)}
    if (alternative =="two.sided"){
        power<-1-pnorm(qnorm(1-sig.level/2)-
                lambda*gamma/sigma)+pnorm(-qnorm(1-
    sig.level/2)- lambda*gamma/sigma)}
    list(power=power)
}
```

(iii) Examples

```
X = c(12.6, 16.2, 13.7, 13.2, 13.7, 13.2, 10.6, 12.7, 15.2, 11.4, 13.8,
14.0, 14.8, 12.9, 12.9, 12.7, 14.2, 12.9 ,13.8,  9.8, 12.8, 11.1, 11.6,
12.9, 13.5, 12.8, 12.2, 14.2, 13.3,  9.8, 12.2, 13.5, 12.6, 12.8, 13.5,
12.1, 14.2, 11.9, 11.3, 10.6, 12.6 ,14.5, 12.5, 15.4, 18.9,  9.9, 12.9,
12.9, 11.0, 12.0, 13.4, 11.3, 15.1, 21.4, 22.2, 20.0)

fit1<-m.test1(X, alternative = "two.sided", mu.not=13.5, sig.level=0.05)
power.m.test1(length(X),alternative="two.sided",delta=1 ,fit1$sigma,
fit1$gamma,0.05)
```

(iv) M-test for testing the equality of location of two populations

```
m.test2<-function(X, Y, alternative = c("two.sided", "less", "greater"),
psi.function =c("psi.huber", "psi.bisquare", "psi.hampel"),sig.level){
    library(MASS)
```

```
    n1<-length(X)
    n2<-length(Y)
    n<-n1+n2
    Z<-c(X,Y)
    ci<-c(rep(0,n1),rep(1,n2))
    vec.unit<-rep(1,n)
    if(psi.function =="psi.huber")
    {
        fit.full<-rlm(matrix(c(vec.unit,ci),ncol=2),Z)
        psi.full<-psi.huber(fit.full$res/mad(fit.full$res),deriv=0)*
                (fit.full$res/mad(fit.full$res))
        sigma.full<-sqrt(sum(psi.full*psi.full)/n)
        fit.null<-rlm(Z~1) #fit.null$s !=mad(Z)

        psi.null<-psi.huber((Z-fit.null$coef)/mad(Z),deriv=0)*
        ((Z-fit.null$coef)/mad(Z))
        sigma.null<-sqrt(sum(psi.null*psi.null)/n)
        gamma <-(1/n)*(sum(psi.huber(fit.full$res/mad(fit.full$res),
        deriv = 1)))/mad(fit.full$res)
}
    if(psi.function =="psi.bisquare")
    {
        fit.full<-rlm(matrix(c(vec.unit,ci),ncol=2),Z, psi=psi.bisquare)
        psi.full<-psi.bisquare(fit.full$res/mad(fit.full$res),deriv=0)*
                    (fit.full$res/mad(fit.full$res))
        sigma.full<-sqrt(sum(psi.full*psi.full)/n)
        fit.null<-rlm(Z~1, psi=psi.bisquare)
        psi.null<-psi.bisquare((Z-fit.null$coef)/mad(Z),deriv=0)*
                    ((Z-fit.null$coef)/mad(Z))
        sigma.null<-sqrt(sum(psi.null*psi.null)/n)
        gamma <-(1/n)*(sum(psi.bisquare(fit.full$res/mad(fit.full$res),
                deriv = 1)))/mad(fit.full$res)
    }
    if(psi.function =="psi.hampel")
    {
        fit.full<-rlm(matrix(c(vec.unit,ci),ncol=2),Z, psi=psi.hampel)
        psi.full<-psi.hampel(fit.full$res/mad(fit.full$res),deriv=0)*
                (fit.full$res/mad(fit.full$res))
        sigma.full<-sqrt(sum(psi.full*psi.full)/n)
        fit.null<-rlm(Z~1, psi=psi.hampel)
        psi.null<-psi.hampel((Z-fit.null$coef)/mad(Z),deriv=0)*
```

```
                    ((Z-fit.null$coef)/mad(Z))
        sigma.null<-sqrt(sum(psi.null*psi.null)/n)
        gamma <-(1/n)*(sum(psi.hampel(fit.full$res/mad(fit.full$res),
                deriv = 1)))/mad(fit.full$res)
    }

    M.stat <-sum(ci*psi.null)/sqrt(n*(sigma.null^2)*n1*n2/(n^2))
    if (alternative =="greater")
    {
        p.value<-1-pnorm(M.stat)
    }
    if  (alternative =="less")
    {
        p.value <-pnorm(M.stat)
    }
    if (alternative =="two.sided")
    {
    p.value <-if (M.stat>=0) 2*(1-pnorm(M.stat)) else 2*pnorm(M.stat)
    }
    list(Mstat=M.stat, p.value=p.value, sigma=sigma.full, gamma=gamma)
    }

(v) The asymptotic power of the M-test for testing location of
two populations

power.m.test2<-function(n1, n, alternative=c("one.sided","two.sided"),
delta, sigma, gamma, sig.level){
    lambda<-delta*sqrt(n)
    if (alternative =="one.sided"){
        power <-1-pnorm(qnorm(1-sig.level)-
            sqrt(n1*(n-n1)/(n^2))*lambda*gamma/sigma)}
    if (alternative =="two.sided"){
        power<-1-pnorm(qnorm(1-sig.level/2)-sqrt(n1*(n-n1)/(n^2))*lambda*
        gamma/sigma)+pnorm(-qnorm(1-sig.level/2)-sqrt(n1*(n-n1)/(n^2))*
        lambda*gamma/sigma)}
    list(power=power)
}

(vi) Examples
iodi.dose<-c(32.5, 40, 46, 43.6, 44.8, 124.5, 40, 33.9, 52, 17, 47.36,
23.4, 34.4, 27.84, 58.43, 38.9, 25.8, 33.6, 61.7, 32.96, 63.7, 56)
```

```
locm.dose<-c(39.4, 40, 56, 49.5, 51.1, 117.6, 40.4, 36.1, 57, 16, 45.84,
24, 35.7, 27.65, 74.22, 46.5, 26.1, 35.3, 68.4, 36.05, 76.4, 60.9)
X<-iodi.dose
Y<-locm.dose
delta<- 2
n1<-length(X)
n<-length(c(X,Y))
fit2<-m.test2(X, Y, alternative = "two.sided", 0.05)
power.m.test2(n1,n, alternative="two.sided", delta, fit2$sigma,
fit2$gamma, 0.05)
```