TESTING FOR OUTLIERS

TESTING FOR OUTLIERS

By

SATWANT CHAWLA, B.A. (HONS.)

A Project
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Master of Science

McMaster University
May 1975

MASTER OF SCIENCE (Statistics)

McMASTER UNIVERSITY Hamilton, Ontario

TITLE: Testing For Outliers

AUTHOR: Satwant Chawla, B.A. (HONS.) [Delhi University]

SUPERVISOR: Professor M. L. Tiku

NUMBER OF PAGES: v, 37

ABSTRACT

Several test statistics, which are known, can be used for testing for outliers. Two new statistics T and t_c are proposed. T and t_c are based on censored and complete samples and are similar to Tiku's T and t_c statistics for testing for normality. The distribution of T is closely approximated by the Beta distribution, and the distribution of t_c is closely approximated by Student's t distribution. T and t_c are also both origin and scale invariant. Besides T and t_c are easy to calculate. The statistic T is more powerful than Tietjen and Moore's statistics L_r and E_r. The statistic t_c is, on the whole, as powerful as E_r.

ACKNOWLEDGEMENTS

The author has pleasure in expressing sincere gratitude to her supervisor, Professor M. L. Tiku for his guidance, valuable suggestions and discussions throughout the course of this study and during the preparation of this project.

The author also wishes to express her sincere appreciation to Dr. J. S. Chawla, her husband, for his encouragement and moral support throughout the duration of this work.

Special thanks are also extended to Mr. Anami Bhargava for his assistance in programming.

Financial assistance by McMaster University is gratefully acknowledged.

Finally, the author would like to thank Miss Linda Westfall for her excellent typing of this manuscript.

TABLE OF CONTENTS

INTRODUCTION

It often happens that in experimental data a few observations occur in a sample which are so far removed from the remaining bulk of observations that the analyst is not willing to believe that these values have come from the same population. Such values are called outliers. The problem of outlier detection has generally been treated as a problem in hypothesis testing. The null hypothesis is that all the observations in a random sample of size n come from the same population $N(\mu,\sigma)$; μ and σ are unknown, against the alternative hypothesis that some specific observations (largest or smallest for example) are outliers, that is, are too large or small as compared to the bulk of observations; see references [9, 10]. The largest observations for example could come from a population with location parameter $\mu + \delta\sigma$, $\delta > 0$, and scale parameter σ. Note that this problem is different than testing that some arbitrary observations are outliers; see Grubbs [9, p. 8]. There is a large number of statistics for testing the largest (smallest) observation in a sample as outliers, but the most prominent statistic is due to Grubbs [9]. This statistic is the ratio of the sum of squares of deviations from the sample mean for a reduced sample (obtained by omitting the largest (smallest) observation) to the sum of squares of deviations for the whole sample. Its generalization for testing r largest (smallest) observations for outliers is due to Grubbs [9, 10] for $r = 2$ and Tietjen and Moore [16] for general r. The Tietjen and Moore statistic L_r, like Grubbs statistic, is the ratio of the sum of squares of deviations for a reduced sample (obtained by omitting the r largest observations) to the sum of squares of deviations for the whole sample. Tietjen and Moore also propose a statistic E_r which can be used for testing r_1 smallest and r_2 largest observations, $r_1 + r_2 = r$. However, the distributions of L_r and E_r are not known; but, Tietjen and Moore provide their Monte Carlo percentage points for values of $n \leq 50$. Sequential tests have been proposed for testing two or more observations and have been based on the statistics due to Grubbs [9], Dixon [4], David et. al [3] and Shapiro and Wilk statistic W [15], also sample skewness $\sqrt{b_1}$ and sample kurtosis b_2; Ferguson [7, 8]. However these sequential tests are known to be ineffective in the presence of masking effect (the phenomenon of some observations being closer to each other than they are close to the bulk of observations); see references [13, 9, 16]. We propose statistics T and t_c for testing r_1 smallest and r_2 largest observations in a sample from a normal population $N(\mu,\sigma)$. T is the ratio of the maximum likelihood estimators of σ calculated from censored and complete samples. The distribution of T under H_0 tends to normality with increasing sample size n (effectively $n > 30$). We propose another statistic, t_c, based on the difference of means of censored and complete samples for testing r_1 smallest and r_2 largest observations. The distribution of t_c is approximately Student's t having n - 1 degrees of freedom. The power of T and t_c is shown to be higher than that of Tietjen and Moore's statistics L_r and E_r.

CHAPTER 1

VARIOUS STATISTICS FOR TESTING SUSPECTED OUTLIERS

To test whether outliers are present in a sample, irrespective of how many, it has been proposed that one of the tests for normality be employed; Feguson [7, 8]. The following tests of normality seem very relevent.

1.1 Various Statistics For Testing Normality

Shapiro And Wilk Statistic.

Shaprio and Wilk [15] proposed a statistic W which is the ratio of the square of an appropriate linear combination of the sample ordered observations to the sum of squares of deviations for the entire sample. Let $m' = (m_1, m_2, \ldots, m_n)$ denote the vector of expected values of standardized normal ordered observations and $V = (v_{ij})$ be the corresponding $n \times n$ variance covariance matrix.

Let $X' = (X_1, X_2, \ldots, X_n)$ denote the vector of ordered observations. If $\{X_i\}$ is an ordered sample from a normal distribution with mean μ and variance σ^2, the best linear unbiased estimator of σ is

$$\sigma' = m'V^{-1}X/m'V^{-1}m \quad .$$

Let

$$S^2 = \sum_1^n (X_i-\bar{X})^2, \ \bar{X} = \frac{1}{n}\sum_1^n X_i.$$

The statistic W is defined

$$W = (a'X)^2/S^2 = (\sum_1^n a_i X_i)^2 / \sum_1^n (X_i - \bar{X})^2 ,$$

where $a' = (a_1, a_2, \ldots, a_n) = m'V^{-1}/(m'V^{-1}m)^{1/2}$. The coefficients $\{a_i\}$ are given in Shapiro and Wilk [15]. The exact distribution of W is not known. The Monte Carlo percentage points of W for $n \leq 50$ are given in [15].

Tiku's T Statistic.

Tiku [21] proposed a statistic T which is the ratio of the estimators of σ caluclated from censored and complete samples.

Let

$$X_1, X_2, \ldots, X_n \tag{1.1}$$

denote the sample of ordered observations. Since the end observations are more sensitive to non-normality, especially to long tailedness, we censor r_1 smallest and r_2 largest observations to obtain the censored sample

$$X_a, X_{a+1}, \ldots, X_b \quad (a=r_1+1, b=n-r_2) \tag{1.2}$$

Let

$$\sigma_c = \{B + \sqrt{B^2+4AC}\}/2A ;$$

be Tiku's modified maximum likelihood estimator of the population standard deviation σ calculated from the censored sample (1.2).

Here

$$A = 1 - q_1 - q_2 \qquad (q_1 = r_1/n, q_2 = r_2/n)$$

$$B = q_2\alpha_2 X_b - q_1\alpha_1 X_a - (q_2\alpha_2 - q_1\alpha_1)K,$$

and

$$C = (\tfrac{1}{n}) \sum_{i=a}^{b} X_i^2 + q_2\beta_2 X_b^2 - q_1\beta_1 X_a^2 - (1-q_1-q_2+q_2\beta_2-q_1\beta_1)K^2 \; ;$$

where

$$K = (\tfrac{1}{n} \sum_{i=a}^{b} X_i + q_2\beta_2 X_b - q_1\beta_1 X_a)/(1-q_1-q_2+q_2\beta_2-q_1\beta_1)$$

$$\beta_1 = -f(t_1)\{t_1+f(t_1)/q_1\}/q_1, \; \alpha_1 = (f(t_1)/q_1) - \beta_1 t_1 \; ,$$

$$\beta_2 = -f(t_2)\{t_2-f(t_2)/q_2\}/q_2, \; \alpha_2 = (f(t_2)/q_2) - \beta_2 t_2 \; ,$$

$$f(t) = (1/\sqrt{2\pi})\exp(-\tfrac{1}{2}\, t^2), \; F(t_1) = \int_{-\infty}^{t_1} f(t)dt = q_1, \; F(t_2) = 1 - q_2.$$

The values of α's and β's can easily be calculated from tables of normal probability function (Biometrika Tables).

Let $\hat{\sigma} = s = \sqrt{\sum_{i=1}^{n} (X_i-\bar{X})^2/n}$ be the maximum likelihood estimator of σ calculated from the complete sample (1.1). The statistic T is defined

$$T = (1-\tfrac{1}{n})\sigma_c/(1-\tfrac{1}{nA})\hat{\sigma}, \; 0 < T < \infty, \; (A=1-q_1-q_2)$$

Small values of T lead to the rejection of normality. The null distribution of T tends to normality with increasing sample size n (effectively n > 30).

Tiku's t_c Statistic.

Tiku [20] proposed another statistic, based on the difference of means of censored and complete samples, for testing normality.

The maximum likelihood estimator of μ (mean of the normal population) calculated from the complete sample (1.1) is $\bar{X} = \sum_{i=1}^{n} X_i/n$. An estimator of μ calculated from the censored sample (1.2) with $r_1 = r_2 = r$ (symmetric censoring) is

$$\mu_c = \{\frac{1}{n}\sum_{i=r+1}^{n-r} X_i + q\beta(X_{r+1}+X_{n-r})\}/d,$$

and

$$t_c = (\mu_c-\bar{X})/s\sqrt{(1-d)/nd},$$

where

$$q = r/n,\ d = 1 - 2q + 2q\beta \text{ and } \beta = -f(t)\{t-f(t)/q\}/q;$$

$$f(t) = \frac{1}{\sqrt{2\pi}}\exp(-\tfrac{1}{2}t^2),\ F(t) = \int_{-\infty}^{t} f(t)dt = 1 - q .$$

The details about t_c are given in Chapter 2.

Other tests for normality are based on

$$\sqrt{b_1} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^3/s^3 \quad \text{(sample skewness)}$$

and

$$b_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^4/s^4 \quad \text{(sample kurtosis)} .$$

Approximations to the distribution of $\sqrt{b_1}$ and b_2 are available in [2]. The exact distributions of $\sqrt{b_1}$ and b_2 are not known.

Another problem is to test for specific number of outliers in an ordered sample, say r_1 on the left and r_2 on the right. The following statistics have been proposed.

1.2 Tests For Detecting One Outlier

Pearson And Chandra Sekar's Statistics.

Pearson and Chandra Sekar [13] criterion is the ratio of the extreme deviate from the sample mean to the estimate of the population standard deviation s

Pearson and Chandra Sekar [13] considered

$$D_n = (X_n-\bar{X})/s, \text{ where } s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

as criteria for rejecting the largest observation, i.e., $r_1 = 0$, $r_2 = 1$, as outlier.

Similarly,

$$D_1 = (\bar{X}-X_1)/s$$

is used as a criteria for rejecting the smallest observation, i.e., $r_1 = 1$, $r_2 = 0$, as outlier.

Grubbs' Statistic.

Grubbs [9] introduced a criteria which is based on the ratio of the sum of squares of deviations for a reduced sample (obtained by omitting the largest (smallest) observation) to the sum of squares of deviations for the whole sample. For testing the largest observation, i.e., $r_1 = 0$, $r_2 = 1$ as outlier, Grubbs [9] proposed the statistic

$$S_n^2/S^2 = \frac{\sum_{i=1}^{n-1}(X_i-\bar{X}_n)^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}\ , \quad \bar{X}_n = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i\ .$$

It turns out that

$$S_n^2/S^2 = 1 - \frac{1}{n-1}\left(\frac{X_n-\bar{X}}{s}\right)^2 = 1 - \frac{1}{n-1} D_n^2$$

where $s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$, and D_n is the studentized maximum residual suggested by E. Pearson and C. Chandra Sekar [13] for testing the significance of the largest observation. If S_n^2/S^2 is smaller than a prescribed percentage points chosen from the tables of Grubbs' paper to achieve a fixed probability level, the largest observation is to be rejected from the data.

Similarly, for testing the smallest observation, i.e., $r_1 = 1$, $r_2 = 0$, as outlier, Grubbs [9] proposed the statistic

$$S_1^2/S^2 = \frac{\sum_{i=2}^{n}(X_i-\bar{X}_1)^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}; \text{ where } \bar{X}_1 = \frac{1}{n-1}\sum_{i=2}^{n} X_i\ .$$

It turns out that

$$S_1^2/S^2 = 1 - \frac{1}{n-1}\left(\frac{\bar{X}-X_1}{s}\right)^2 = 1 - \frac{1}{n-1} D_1^2$$

where $s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$ and D_1 is the studentized minimum residual. If S_1^2/S^2 is too small, the smallest value is to be rejected. The statistic D_n (or D_1) is easier to compute than S_n^2/S^2 (or S_1^2/S^2). Grubbs [10] recommended that if one is interested only in outliers that occur on the right hand side or high side, one should use the statistic

$$D_n = (X_n - \bar{X})/s \ .$$

On the other hand, if one is interested only in outliers occuring on the left hand side or low side, one should always use the statistic

$$D_1 = (\bar{X} - X_1)/s \ .$$

The percentage points of D_n and D_1 are given in Table 1 of Grubbs [10].

Dixon's Statistics.

Dixon [4] criteria are ratios of the distance between a suspected observation and its nearest or next nearest (assumed unaffected) neighbor to the range of the sample (with zero, one, or two observations omitted). Dixon's test is based on a value being too large (small) compared with its nearest neighbor.

For testing the largest observations, i.e., $r_1 = 0$, $r_2 = 1$, as outlier, Dixon [4, 6] proposed the following statistics

$$r_{ij} = \frac{X_n - X_{(n-i)}}{X_n - X_{(j+1)}} \ , \ i = 1, 2; \ \ j = 0, 1, 2 \ . \qquad (1.5)$$

If (1.5) is larger than a specified percentage point found in the tables of Dixon's paper [6], the largest observation is to be rejected from the data. Dixon [5] also proposed the statistics

$$r_{ij} = \frac{X_{i+1}-X_1}{X_{n-j}-X_1}\ ,\ i = 1,\ 2;\quad j = 0,\ 1,\ 2 \tag{1.6}$$

for testing the smallest observation i.e., $r_1 = 1$, $r_2 = 0$ as outlier. If $|r_{ij}|$ (r_{ij} given by (1.6)) is too large, the smallest observation is to be rejected from the data.

The two sided-extension of Dixon's rule, suggested by Ferguson [8] for the rejection of one outlier would be to reject the largest or smallest observation according to whether X_n-X_{n-1} is larger or smaller than X_2-X_1, whenever

$$\max\left(\frac{X_n-X_{n-1}}{X_n-X_1}\ ,\ \frac{X_2-X_1}{X_n-X_1}\right)$$

is too large. Tables for the exact percentage points for this rule are not available; however as King [12] points out, approximate percentage points can be obtained from tables of the one sided tests for doubling the significance level.

1.3 Tests For Detecting Two Outliers

Grubbs' Statistic.

For testing the two largest (or smallest) observations as outliers, Grubbs [9] also proposed the statistics:

$$S^2_{n-1,n}/S^2 = \frac{\sum_{i=1}^{n-2}(X_i-\bar{X}_{n-1,n})^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}, \text{ where } \bar{X}_{n-1,n} = \frac{1}{n-2}\sum_{i=1}^{n-2}X_i$$

for testing two largest observations, i.e., $r_1 = 0$, $r_2 = 2$, as outliers; and

$$S^2_{1,2}/S^2 = \frac{\sum_{i=3}^{n}(X_i-\bar{X}_{1,2})^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}, \quad \bar{X}_{1,2} = \frac{1}{n-2}\sum_{i=3}^{n}X_i$$

for testing two smallest observations i.e., $r_1 = 2$, $r_2 = 0$, as outliers.

David, Pearson And Hartley's Statistic.

David, Pearson and Hartley [3] criterion is the ratio of the sample range to the sample standard deviation. They have proposed the statistic

$$w/s = \frac{X_n-X_1}{s}, \text{ where } s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

for testing the smallest and the largest observations simultaneously (i.e., $r_1 = 1$, $r_2 = 1$) as probable outliers in a sample. If X_n is about as far above the mean, $\bar{X}$, as X_1 is below $\bar{X}$, and if w/s exceeds some chosen critical value, then one would conclude that both the suspected values are outliers. If, however, X_1 and X_n are displaced from the mean by different amounts, some further test would have to be made to decide whether to reject as outlying only the lowest or only the highest value or both the lowest and highest values.

1.4 Tests For Detecting Several Outliers

Tietjen And Moore's Statistics.

Tietjen and Moore [16] have proposed the statistic:

$$L_r = \sum_{i=1}^{n-r}(X_i-\bar{X}_r)^2/\sum_{i=1}^{n}(X_i-\bar{X})^2,\ \bar{X}_r = \sum_{i=1}^{n-r}X_i/n-r\ ,$$

for testing the largest r observations i.e., $r_1 = 0$, $r_2 = r$, as outliers, and

$$L_r^* = \sum_{i=r+1}^{n}(X_i-\bar{X}_r^*)^2/\sum_{i=1}^{n}(X_i-\bar{X})^2,\ \text{where } \bar{X}_r^* = \frac{1}{n-r}\sum_{i=r+1}^{n}X_i$$

for testing the smallest r observations, i.e., $r_1 = r$, $r_2 = 0$, as outliers.

Note that L_1 is equal to Grubbs' S_n^2/S^2 and L_2 is equal to Grubbs' $S_{n-1,n}^2/S^2$. Similarly, it should be noted that L_1^* is Grubbs' S_1^2/S^2 and L_2^* is Grubbs' $S_{1,2}^2/S^2$.

The L_r and L_r^* statistics are useful for examining r suspected values which are larger or smaller than the bulk of the sample. Some samples, however, display suspect values on both sides of the bulk. To deal with the situation in which some of the r suspected values are larger and some are smaller than the remaining values, another statistic is suggested by Tietjen and Moore. Again let the sample values be denoted by $x_1, \ldots, x_n$. Compute the mean of the sample, $\bar{x}$. Then compute the n absolute residuals:

$$R_1 = |x_1-\bar{x}|,\ R_2 = |x_2-\bar{x}|,\ \ldots,\ R_n = |x_n-\bar{x}|\ .$$

Now relable the observations as Y's in such a manner that Y_i is the x whose R_i is the ith largest. This means that Y_1 is the observation closest to the mean $\bar{x}$ and that Y_n is the observation farthest from the mean. The proposed statistic for testing r_1 smallest and r_2 largest observations $(r_1 + r_2 = r)$ is

$$E_r = \sum_{i=1}^{n-r}(Y_i-\bar{Y}_r)^2 / \sum_{i=1}^{n}(Y_i-\bar{Y})^2$$

where

$$\bar{Y}_r = \sum_{i=1}^{n-r} Y_i/(n-r)$$

is the mean of the (n-r) least extreme observations and $\bar{Y}$ is the mean of the full sample. The Monte Carlo percentage points of $L_r(L_r^*)$ and E_r are given in [16].

The statistics $L_r(L_r^*)$ and E_r are used as follows: if in a sample of size n, we decide to test whether the r largest (smallest) values are outliers, we calculate $L_r(L_r^*)$, if this quantity is smaller than the desired critical value in Table I of Tietjen and Moore [16], we conclude that the r largest (smallest) values are indeed outliers. If on the other hand we wish to check whether the r "most extreme" (as measured from the sample mean) values are outliers, we calculate E_r; if this quantity is smaller than the selected critical value in Table II of Tietjen and Moore [16], we conclude that the r suspected observations are outliers.

$\sqrt{b_1}$ and b_2 given by Ferguson [7, 8] (defined as earlier) do not depend on r_1 and r_2 and may therefore be also used for testing specified number of outliers. Particularly, $\sqrt{b_1}$ is suggested for use in the one-sided situation (change in level of several observations in the same direction), for example where it is known that all the possible spurious observations are too large. The coefficient of kurtosis b_2 is suggested for use in the "two sided" tests (change in level to higher and lower values) and also for changes in scale (variance).

If $\sqrt{b_1}$ is too large the largest observation is to be rejected as spurious; and the same procedure repeated until no further sample values are judged as outliers. If b_2 is too large, the observation farthest from the mean is to be rejected as spurious. Similarly this kurtosis test can be used sequentially. However, any such sequential procedure is ineffective in the presence of "masking effect".

In practice r_1 and r_2 in the above statistics might not be known. In such situations, one-at-a-time sequential tests proposed by Ferguson [7, 8], Dixon [4, 5, 6], David et.al [3] and Shapiro and Wilk [15] can be used. But the difficulty in applying these tests is that, if we have a few suspected observations far away from the bulk but closer to each other, the above tests will not reject either one no matter how suspected they may be in appearance. This is what Pearson and Chandra Sekar [13] refer to as the masking effect, namely, that the presence of several suspected observations reduces the ability of the rejection rule to detect even one.

CHAPTER 2

DISTRIBUTIONS AND POWERS OF THE STATISTICS T AND t_c

In this chapter the statistics T and t_c are used for detecting several outliers in a sample of size n from a normal population $N(\mu,\sigma)$; μ and σ are unknown. The powers of these statistics are calculated for various sample sizes and it is shown that the power of these statistics is higher than that of Tietjen and Moore's statistics L_r and E_r in detecting shifts in location.

2.1 Statistic T for Testing Suspected Outliers

Let

$$x_1, x_2, \ldots, x_n \tag{2.1.1}$$

be a random sample from a normal population $N(\mu,\sigma)$, μ and σ are unknown. We order the observations according to increasing magnitude and denote the ith largest observation by X_i; thus

$$X_1 \leq X_2 \leq \ldots \leq X_n \tag{2.1.2}$$

is the ordered set of observations.

The most commonly encountered situation is to test whether some particular observations (r_2 largest and r_1 smallest observations for example) are too large or small as compared to the remaining bulk of observations. To develop a statistic for testing r_1 smallest and

r_2 largest observations in (2.1.2), we consider the censored sample

$$X_a,\ X_{a+1},\ \ldots,\ X_b \quad (a=r_1+1,\ b=n-r_2)\ . \qquad (2.1.3)$$

Let σ_c be the maximum likelihood estimator, or an estimator which is identical, at least asymptotically, to the maximum likelihood estimator (for example Tiku's [17] modified maximum likelihood estimator) of the population standard deviation σ calculated from the censored sample (2.1.3) and let $\hat{\sigma}$ be the maximum likelihood estimator of σ calculated from the complete sample (2.1.1) or (2.1.2). Consider the statistic

$$T = \frac{(1-\frac{1}{n})\sigma_c}{(1-\frac{1}{nA})\hat{\sigma}}\ ,\quad 0 < T < \infty \quad (A=1-q_1-q_2) \qquad (2.1.4)$$

the expressions for σ_c and $\hat{\sigma}$ are given in Chapter 1. Note that T is origin and scale invariant. The statistic T is proposed as a test-statistic for testing the null hypothesis

H_0: that the sample contains no outlier, i.e., that all the observations in (2.1.1) come from the same population $N(\mu,\sigma)$;

against the alternative hypothesis

H_1: that r_1 smallest and r_2 largest observations are outliers, i.e., are too small and too large, respectively, as compared to the bulk of observations.

Small values of T lead to the rejection of H_0.

For fixed $q_1 = r_1/n$ and $q_2 = r_2/n$, the asymptotic null distribution of T is normal. This is because of the fact that for fixed q_1 and q_2, $\hat{\sigma}$ converges to σ faster than σ_c and since σ_c is the maximum likelihood estimator (or identical to it asymptotically), the asymptotic distribution of $\sigma_c/\hat{\sigma} \approx \sigma_c/\sigma$ and hence of T is normal with E(T) = 1 and

$$V(T) \approx \{(1-\frac{1}{n})/(1-\frac{1}{nA})\}^2 \cdot (1/\sigma^2)\{V(\sigma_c)+V(\hat{\sigma})-2\,\mathrm{cov}(\sigma_c,\hat{\sigma})\} \qquad (2.1.5)$$

This is because (Kendal and Stuart [11])

$$V(x/y) \approx \{\frac{E(x)}{E(y)}\}^2 \{\frac{V(x)}{E^2(x)} + \frac{V(y)}{E^2(y)} - 2\,\frac{Cov(x,y)}{E(x)E(y)}\},$$

and σ_c and $\hat{\sigma}$ are asymptotically unbiased. It can be shown that (Tiku [22])

$$(n/\sigma^2)V(\sigma_c) \approx 1/\{2(1-q_1-q_2)-(q_2\alpha_2 t_2-q_1\alpha_1 t_1)\}$$
$$(n/\sigma^2)V(\hat{\sigma}) \approx \tfrac{1}{2} \quad \text{and}$$
$$(n/\sigma^2)Cov(\sigma_c,\hat{\sigma}) \approx \tfrac{1}{2}\text{ , for small } q_1 \text{ and } q_2\text{, where}$$

$$P(t_1) = q_1 = \int_{-\infty}^{t_1} f(z)\,dz,\ P(t_2) = 1-q_2,\ f(z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2}.$$

Tiku [21] used the statistic T with $r_1 = 0$ and $r_2 = [\frac{1}{2}+0.6n]$ as a test for normality against positively skewed distributions and with $r_1 = r_2 = [\frac{1}{2}+0.3n]$ as a test for normality against symmetric distributions. For these values of r_1 and r_2, the distribution of T is effectively normal for $n \geq 30$; see Tiku [21, 22]. $[f]$ denotes the integer value of f. For small values of r_1 and r_2, Tiku [23] suggests that the 100α percentage points of T may be obtained from the following Beta-approximation

$$\{(n-1)/(n-r_1-r_2-1)\}u_\alpha + (1/5n)\{1+1/(n-2r_{2+1})\} ,$$

where u_α is the 100α percentage point of Beta distribution $\beta(n-r_1-r_2-1, r_1+r_2)$. An extensive comparison of these approximate values with the Monte Carlo values is given in Table 1.

2.2 Percentage Points of T

Samples of size n = 8, 10, 12, 16, 20, 24, 30 from N(0,1) were generated and the lower 100α percentage points of T simulated; α = .01, .05, .10. These values are given in Table 1. Also given are the corresponding values calculated from the above Beta Approximation. It is clear that the agreement between the two is very good.

Note that the distribution of T with $r_1 = i$ and $r_2 = j$ is the same as of T with $r_1 = j$ and $r_2 = i$, because normal distribution is symmetric. We therefore only considered $r_1 \leq r_2$.

2.3 Statistic t_c For Testing Suspected Outliers

We propose t_c as a test statistic for testing the null hypothesis H_0 against the alternative hypothesis H_1; where H_0 and

H_1 are as given in section 2.1. Let

$$r = \max(r_1, r_2).$$

Define the statistic

$$t_c = (\mu_c - \bar{x})/s\sqrt{(1-d)/nd}\ , \quad d = 1-2q+2q\beta \qquad (2.3.1)$$

where

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2, \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i = \frac{1}{n}\sum_{i=1}^{n}X_i = \bar{X}$$

and

$$\mu_c = \{\frac{1}{n}\sum_{i=r+1}^{n-r}X_i + q\beta(X_{r+1}+X_{n-r})\}/d \qquad (2.3.2)$$

is modified maximum likelihood estimator of μ calculated from the symmetrically censored sample (see Tiku [19, 20])

$$X_{r+1}, X_{r+2}, \ldots, X_{n-r}\ .$$

Here

$$\beta = -f(t)\{t-f(t)/q\}/q\ ,$$

$$f(t) = \frac{1}{\sqrt{2\pi}}\ \exp(-\frac{1}{2}\ t^2) \text{ and } F(t) = \int_{-\infty}^{t} f(t)dt = 1-q\ .$$

Note that t_c is both origin and scale invariant.
Since for normal samples μ_c is the solution of the maximum likelihood equation (see Tiku [17, 19])

$$d\ \log L/d\mu = nd(\mu_c-\mu)/\sigma^2$$

for which $d^r \log L/d\mu^r$, $r \geq 3$, are zero, it follows from Bartlett [1] that for large n, μ_c is normally distributed with mean μ and variance

$$V(\mu_c) = 1/-E(d^2 \log L/d\mu^2) = \sigma^2/nd\ .$$

In fact the distribution of μ_c tends to normality very rapidly; see Tiku [20] .

Since in independent random sampling $\mu_c-\bar{x}$ assumes independent values, and $\bar{x}$ is also normally distributed, the distribution of $\mu_c-\bar{x}$ is normal with mean zero and variance

$$V(\mu_c-\bar{x}) = \sigma^2(1-d)/nd\ . \qquad (2.3.3)$$

Equation (2.3.3) is true because $\text{cov}(\mu_c,\bar{x}) = \sigma^2/n$, since $\text{cov}(X_i,\bar{x}) = \sigma^2/n$.

Since in case of symmetric censoring, the numerator in (2.3.1) is of the form $\sum_{i=1}^{n} l_i(X_i+X_{n-i+1})$, it is uncorrelated with s. Therefore, the distribution of t_c is approximately t having n-1 degrees of freedom. In fact, student's t distribution provides very accurate values of the percentage points of t_c, even for small samples. For example for n = 10, the empirical values (based on 20,000 samples) of the 5% and 10% significance levels of t_c are 2.27 and 1.84 respectively. The corresponding values of student's t distribution having 9 degrees of freedom are 2.26 and 1.84. Large positive and negative values of t_c lead to the rejection of H_0.

2.4 Power of T and t_c Under Grubbs Model

There may be many models for outliers, but an important practical case involves the situation where all the observations in the sample have the same standard error, whereas the largest or smallest observations may result from shifts in level. For example, if the largest observations appear unusually high compared to other observations in the sample, we may want to consider the model

Model A

H_0: that all the observations in (2.1.1) come from $N(\mu,\sigma)$; μ and σ are unknown, against

H_1: that r_2 largest observations in (2.1.2) come from $N(\mu+\delta\sigma,\sigma)$, $\delta > 0$.

Another case involves the situation where the largest and/or smallest observations are subject to different standard errors. For example if the largest observations appear unusually high we may consider the model (see Grubbs [9, p. 41])

Model B

H_0: that all the observations come from $N(\mu,\sigma)$, against

H_1: that r_2 largest observations come from $N(\mu,\delta\sigma)$, $\delta > 1$.

Samples of size n were taken from N(0,1), and to a fixed member of each sample was added, successively, the non-negative numbers $\delta = 0.5, 1.0, 1.5, 2.0, 3.0$. The values of n chosen for this study were n = 10, 20, 30. The empirical values (based on 2000 random samples) of the power of T under Model A are given in Table II and are compared with the values of the power of Tietjen and Moore's statistic L_r. It is clear from table II that T is more powerful than L_r.

We also calculated the power of T against two-sided models

H_1: X_i come from $N(\mu-\delta_1\sigma,\sigma)$ for $i = 1, 2, \ldots, r_1$ and

X_i come from $N(\mu+\delta_2\sigma,\sigma)$ for $i = n-r_2+1, \ldots, n$,

$\delta_1 > 0$, $\delta_2 > 0$

and compared with the values of the power of $E_r(r=r_1+r_2)$. The values of the power of T, E_r and t_c are given in Table III. It is clear from tables II and III that the statistic T is more powerful than L_r and E_r. It is ~~also~~ clear from these table (table III) that the statistic t_c has no such power-superiority over the statistic E_r.

EXAMPLE 1

We take this example from Tietjen and Moore [16, p. 590]. A solution was analyzed for a single isotope of uranium by mass spectrometry methods. Eight observations on the sample, arranged in ascending order, are as follows:

.00229, .00236, .00323, .00357, .00363, .00381, .00401, .00408 .

The two smallest observations are highly suspect. It has been shown in Tietjen and Moore [16] that if one uses sequential one-at-a-time procedures, then Grubbs' statistic S_1^2/S^2, Feguson's statistic $\sqrt{b_1}$, and Dixon's statistic r_{11} do not reject the first observation .00229 at 10% significance level, and this is because of the masking effect of the second observation. However, they have shown that the Grubbs statistic $S_{1,2}^2/S^2$ or L_2 do reject the two lowest observations .00229, .00236 simultaneously even at 5% significance level. In what follows, we will show that the statistic T also rejects the observations .00229, .00236 simultaneously at 5% significance level. Here

$$n = 8, r_1 = 2, r_2 = 0, q_1 = .25, q_2 = 0,$$
$$\alpha_1 = 0, \alpha_2 = .7066, \beta_1 = -1.0, \beta_2 = .8171$$
$$\sigma_c = .000373, \hat{\sigma} = .000054$$
$$T = (1-1/8)(.000373)/(1-1/6)(.000654) = .606$$

which is smaller than the 5% critical value of .618. Thus the statistic T rejects the two lowest observations .00229 and .00236 simultaneously at 5% significance level.

EXAMPLE 2

The following example was taken by Grubbs [9, p. 8] .

-1.40	-.44	-.30	-.24	-.22
-.13	-.05	.06	.10	.18
.20	.39	.48	.63	1.01

Using the sequential one-at-a-time procedures, he showed that the statistic S_1^2/S^2 rejects the observation -1.40 as an outlier, but not 1.01. However, the statistic T rejects (and so does Tietjen and Moore's statistic E_2) the two lowest and highest observations -1.40 and 1.01, simultaneously, at 5% significance level.

For T we have,

$n = 15$, $r_1 = r_2 = 1$, $q_1 = q_2 = 0.0667$, $\alpha_1 = \alpha_2 = 0.6407$,
$\beta_2 = -\beta_1 = 0.8711$, $\hat{\sigma} = 0.53226$, and
$\sigma_c = \{.045726 + \sqrt{(.002091+.402747)}\}/1.7332 = 0.39349$;
$T = (1-1/15)(0.39349)/(1-1/12.999)(0.53226) = .747$

which is smaller than the 5% critical value of .813 (Interpolated from Table I). Thus the statistic T rejects the two extreme observations -1.40 and 1.01 simultaneously at 5% significance level.

We also work out this problem by using the Statistic t_c. Here

$$\mu_c = .056002,\ \bar{x} = .018,\ s = .55095 \text{ and } d = .982805;$$
$$t_c = (.056002-.018)/.55095\times\sqrt{(.017195)/14.742075} = 2.01977$$

which is larger than the 10% critical value of 1.76 (for ν=14 d.f.) so that the extreme observations -1.40 and 1.01 would be rejected simulataneously as outliers.

EXAMPLE 3

This example is also taken from Tietjen and Moore [16, p. 593]. A set of eight mass spectrometer measurements on a particular isotope of uranium (different from that of Example 1) is arranged in increasing order of magnitude as

199.31, 199.53, 200.19, 200.82, 201.92, 201.95, 202.18, 245.57 .

If we apply the Statistic T for testing the largest observation, i.e., $r_1 = 0$ and $r_2 = 1$, we get

$$\sigma_c = \{(.1058)+\sqrt{(.01119+5.08896)}\}/1.75 = 1.35094$$

(Note that $\alpha_1 = 0$, $\alpha_2 = .7066$ and $\beta_1 = -1.0$, $\beta_2 = .8171$).

Also $\hat{\sigma} = 14.82872$ and hence $T = (1-1/8)(1.35094)/(1-1/7)(14.82872) = .09300$ which is smaller than the 1% critical value of .576, so that the statistic rejects 245.57 as an outlier. If, however, we apply the same test for testing the two largest observations, i.e., $r_1 = 0$ and $r_2 = 2$, we have $\sigma_c = \{.20145+\sqrt{.04058+3.27879}\}/1.5 = 1.34891$.

Note that in this case $\alpha_1 = 0$, $\alpha_2 = .7595$; $\beta_1 = -1.0$ and $\beta_2 = .7592$. Also $\hat{\sigma} = 14.82872$.

Thus

$$T = (1-1/8)(1.34891)/(1-1/6)(14.82872) = .0955$$

which is smaller than the 1% critical value of .448. So that the statistic T also rejects both 202.18 and 245.57 as outliers. It is clear, however, that 202.18 is not an outlier! This example shows how important it is to use the appropriate value of r_1 and r_2 for T as well as for all the above statistics proposed for testing outliers.

2.5 Choice of r_1 and r_2

From the above examples, it is clear that the choice of r_1 and r_2 is subjective and requires judgement on the part of the user after the data are taken. Since no outliers are anticipated, how many outliers should be tested for? Example 1 illustrates how the test for one outlier may be inappropriate because of the masking effect. Example 3 illustrates how an error can be made in testing for two large observations when only one outlier is actually present. Evidently, the proper choice of r_1 and r_2 is important. In practice, the choice of r_1 and r_2 will have to be made subjectively after the experiment is done and data taken.

TABLE I

CRITICAL VALUES FOR T

$r_1 = 1, r_2 = 1$

r_1	r_2	α = 0.01 Emp:	Approx:	α = 0.05 Emp:	Approx:	α = 0.10 Emp:	Approx:
				n = 8			
0	1	.576	.570	.731	.737	.818	.823
	2	.448	.442	.618	.615	.718	.716
	3	.340	.336	.514	.507	.622	.616
	4	.226	.248	.396	.407	.512	.519
1	1	.432	.440	.602	.614	.704	.715
	2	.350	.333	.504	.504	.610	.613
	3	.216	.232	.372	.390	.486	.502
2	2	.220	.228	.380	.387	.486	.499
				n = 10			
0	1	.653	.655	.795	.796	.865	.866
	2	.551	.550	.703	.704	.787	.787
	3	.475	.463	.637	.624	.724	.717
	4	.392	.383	.552	.548	.648	.648
	5	.305	.312	.470	.474	.580	.580
1	1	.555	.549	.702	.703	.784	.786
	2	.464	.464	.624	.624	.711	.716
	3	.382	.380	.546	.544	.642	.645
	4	.285	.299	.456	.461	.568	.567
2	2	.380	.379	.540	.544	.644	.644
	3	.287	.296	.457	.458	.559	.564
				n = 12			
0	1	.723	.778	.839	.833	.891	.892
	2	.638	.625	.766	.759	.832	.829
	3	.573	.553	.709	.697	.781	.775
	4	.517	.487	.655	.638	.733	.724
	5	.438	.422	.590	.580	.682	.671
	6	.372	.363	.532	.522	.628	.621
1	1	.626	.624	.758	.759	.824	.828
	2	.554	.552	.702	.696	.780	.775
	3	.494	.486	.640	.637	.728	.723
	4	.432	.420	.584	.577	.669	.669
	5	.350	.352	.518	.511	.620	.610

(continued)

TABLE I

CRITICAL VALUES FOR T

		α = 0.01		α = 0.05		α = 0.10	
r_1	r_2	Emp:	Approx:	Emp:	Approx:	Emp:	Approx:
				n = 12			
2	2	.491	.485	.641	.636	.727	.722
	3	.421	.419	.585	.576	.677	.668
	4	.345	.350	.502	.508	.603	.607
3	3	.346	.349	.517	.507	.609	.606
				n = 16			
0	1	.793	.784	.880	.878	.919	.922
	2	.734	.719	.832	.825	.880	.878
	3	.687	.666	.795	.783	.849	.843
	4	.621	.618	.745	.742	.813	.809
	5	.582	.574	.712	.704	.784	.776
	6	.539	.528	.685	.665	.762	.743
	7	.496	.485	.648	.628	.728	.713
	8	.454	.445	.604	.593	.690	.681
1	1	.721	.718	.830	.824	.878	.877
	2	.675	.666	.788	.782	.844	.842
	3	.632	.618	.755	.742	.814	.809
	4	.580	.573	.715	.704	.785	.776
	5	.546	.527	.683	.664	.754	.742
	6	.506	.484	.640	.626	.718	.711
	7	.447	.437	.593	.585	.676	.673
2	2	.630	.618	.750	.742	.811	.809
	3	.593	.573	.717	.704	.783	.776
	4	.542	.527	.668	.664	.752	.742
	5	.498	.483	.635	.625	.715	.710
	6	.433	.435	.591	.583	.672	.671
3	3	.549	.527	.668	.664	.744	.742
	4	.491	.483	.632	.625	.713	.710
	5	.438	.434	.583	.582	.668	.670
4	4	.433	.434	.585	.582	.672	.670

(continued)

TABLE I

CRITICAL VALUES FOR T

r_1	r_2	α = 0.01 Emp:	α = 0.01 Approx:	α = 0.05 Emp:	α = 0.05 Approx:	α = 0.10 Emp:	α = 0.10 Approx:
				n = 20			
0	1	.834	.824	.906	.904	.942	.939
	2	.790	.775	.867	.863	.906	.906
	3	.740	.731	.836	.827	.880	.876
	4	.712	.698	.814	.801	.864	.855
	5	.692	.662	.782	.772	.841	.831
	6	.639	.629	.761	.745	.820	.808
	7	.616	.597	.733	.719	.796	.785
	8	.589	.565	.718	.691	.779	.763
	9	.553	.532	.683	.665	.755	.741
	10	.516	.497	.657	.634	.729	.715
1	1	.783	.774	.867	.862	.907	.905
	2	.742	.731	.833	.827	.876	.876
	3	.710	.698	.808	.800	.852	.855
	4	.676	.662	.783	.772	.840	.831
	5	.636	.629	.754	.750	.813	.808
	6	.619	.597	.735	.719	.795	.785
	7	.573	.564	.698	.690	.762	.762
	8	.542	.531	.678	.664	.749	.740
	9	.504	.490	.634	.627	.710	.708
2	2	.713	.698	.810	.801	.860	.855
	3	.688	.662	.788	.772	.842	.831
	4	.648	.629	.751	.745	.814	.808
	5	.611	.597	.731	.719	.791	.785
	6	.581	.564	.692	.690	.766	.762
	7	.547	.530	.680	.663	.753	.739
	8	.502	.489	.638	.626	.713	.707
3	3	.645	.629	.746	.745	.809	.808
	4	.625	.597	.722	.719	.787	.785
	5	.565	.564	.701	.690	.764	.762
	6	.536	.530	.669	.663	.740	.739
	7	.504	.488	.640	.625	.716	.706
4	4	.573	.564	.692	.690	.761	.762
	5	.521	.530	.663	.663	.737	.739
	6	.495	.488	.641	.626	.719	.706
5	5	.505	.488	.635	.626	.710	.706

(continued)

TABLE I

CRITICAL VALUES FOR T

		α = 0.01		α = 0.05		α = 0.10	
r_1	r_2	Emp:	Approx:	Emp:	Approx:	Emp:	Approx:
				n = 24			
0	1	.861	.852	.923	.918	.949	.947
	2	.824	.809	.890	.884	.924	.920
	3	.796	.779	.866	.861	.906	.901
	4	.767	.746	.843	.833	.886	.879
	5	.741	.717	.830	.810	.876	.859
	6	.724	.690	.810	.788	.861	.841
	7	.694	.667	.792	.771	.842	.827
	8	.662	.645	.767	.754	.824	.813
	9	.636	.621	.753	.734	.810	.798
	10	.618	.590	.725	.709	.789	.774
	11	.602	.565	.714	.688	.778	.759
	12	.559	.539	.677	.667	.749	.740
1	1	.810	.809	.884	.884	.922	.920
	2	.788	.779	.867	.861	.905	.901
	3	.751	.746	.849	.833	.890	.879
	4	.740	.717	.827	.810	.869	.858
	5	.732	.689	.812	.788	.855	.841
	6	.686	.667	.780	.771	.836	.827
	7	.668	.645	.766	.754	.828	.813
	8	.634	.620	.744	.734	.802	.798
	9	.618	.590	.729	.708	.786	.774
	10	.588	.564	.700	.687	.772	.759
	11	.564	.534	.677	.661	.746	.735
2	2	.764	.746	.843	.833	.884	.879
	3	.746	.717	.828	.810	.876	.858
	4	.710	.689	.806	.788	.856	.841
	5	.676	.667	.774	.771	.832	.827
	6	.664	.645	.760	.754	.816	.813
	7	.639	.620	.746	.734	.803	.798
	8	.607	.590	.719	.708	.787	.774
	9	.580	.563	.699	.686	.764	.759
	10	.556	.533	.676	.661	.744	.735
3	3	.712	.689	.803	.810	.852	.841
	4	.680	.667	.776	.771	.827	.827
	5	.656	.645	.771	.754	.823	.813
	6	.640	.620	.745	.734	.800	.798
	7	.596	.589	.718	.708	.783	.774

(continued)

TABLE I

CRITICAL VALUES FOR T

r_1	r_2	α = 0.01 Emp:	α = 0.01 Approx:	α = 0.05 Emp:	α = 0.05 Approx:	α = 0.10 Emp:	α = 0.10 Approx:
				n = 24			
3	8	.578	.563	.691	.686	.762	.759
	9	.551	.532	.674	.661	.747	.735
4	4	.660	.645	.768	.754	.824	.812
	5	.624	.620	.731	.733	.795	.797
	6	.600	.589	.722	.708	.784	.773
	7	.576	.563	.691	.686	.767	.757
	8	.552	.532	.672	.659	.748	.733
5	5	.608	.589	.717	.708	.776	.773
	6	.580	.563	.706	.686	.762	.756
	7	.548	.532	.681	.659	.752	.732
6	6	.544	.532	.665	.659	.741	.732
				n = 30			
0	1	.881	.882	.937	.935	.959	.959
	2	.855	.845	.911	.906	.939	.935
	3	.823	.816	.897	.884	.928	.917
	4	.810	.792	.881	.864	.916	.901
	5	.784	.769	.866	.847	.900	.887
	6	.777	.750	.849	.832	.892	.873
	7	.767	.733	.838	.819	.879	.864
	8	.739	.718	.825	.808	.865	.852
	9	.725	.703	.815	.797	.860	.846
	10	.734	.679	.796	.776	.843	.830
	11	.693	.660	.786	.759	.838	.816
	12	.667	.639	.768	.745	.820	.801
	13	.654	.622	.760	.733	.810	.793
	14	.626	.606	.736	.719	.798	.782
	15	.623	.587	.724	.699	.781	.769
1	1	.858	.850	.913	.906	.942	.935
	2	.837	.816	.896	.884	.924	.917
	3	.803	.792	.877	.864	.911	.901
	4	.789	.769	.861	.847	.897	.887
	5	.767	.750	.845	.832	.884	.873
	6	.755	.733	.836	.819	.880	.864
	7	.747	.718	.820	.808	.862	.852
	8	.721	.703	.810	.797	.858	.846

(continued)

TABLE I

CRITICAL VALUES FOR T

r_1	r_2	α = 0.01 Emp:	α = 0.01 Approx:	α = 0.05 Emp:	α = 0.05 Approx:	α = 0.10 Emp:	α = 0.10 Approx:
				n = 30			
1	9	.694	.679	.791	.776	.844	.830
	10	.688	.660	.784	.759	.835	.816
	11	.664	.638	.763	.744	.818	.800
	12	.659	.622	.754	.733	.806	793
	13	.613	.605	.738	.718	.796	.781
	14	.608	.583	.714	.695	.776	.765
2	2	.808	.792	.878	.864	.910	.901
	3	.793	.769	.868	.847	.902	.887
	4	.761	.750	.847	.832	.887	.873
	5	.747	.733	.832	.819	.875	.864
	6	.742	.718	.815	.808	.862	.852
	7	.721	.703	.812	.797	.854	.846
	8	.696	.679	.789	.776	.836	.830
	9	.678	.660	.783	.759	.829	.816
	10	.669	.638	.764	.744	.819	.800
	11	.641	.621	.747	.732	.805	.792
	12	.629	.605	.732	.718	.789	.780
	13	.608	.582	.717	.694	.779	.764
3	3	.770	.750	.846	.832	.885	.873
	4	.755	.733	.830	.819	.872	.864
	5	.739	.718	.819	.808	.865	.852
	6	.722	.703	.815	.797	.857	.846
	7	.698	.679	.789	.776	.837	.830
	8	.672	.660	.771	.759	.827	.816
	9	.668	.638	.767	.744	.817	.800
	10	.642	.621	.747	.732	.804	.792
	11	.629	.605	.724	.718	.785	.780
	12	.593	.582	.709	.694	.771	.764
4	4	.725	.718	.813	.808	.863	.852
	5	.710	.703	.803	.797	.851	.846
	6	.706	.679	.787	.776	.843	.830
	7	.690	.660	.777	.759	.827	.816
	8	.668	.638	.761	.744	.812	.800
	9	.636	.621	.753	.732	.808	.792
	10	.604	.605	.714	.718	.786	.780
	11	.600	.581	.710	.693	.769	.763

(continued)

TABLE I

CRITICAL VALUES FOR T

r_1	r_2	$\alpha = 0.01$ Emp:	Approx:	$\alpha = 0.05$ Emp:	Approx:	$\alpha = 0.10$ Emp:	Approx:
				n = 30			
5	5	.696	.679	.794	.776	.847	.830
	6	.676	.660	.778	.759	.829	.816
	7	.660	.638	.757	.744	.811	.800
	8	.635	.621	.744	.732	.799	.792
	9	.632	.604	.738	.717	.795	.780
	10	.599	.581	.697	.693	.764	.763
6	6	.652	.638	.757	.738	.814	.800
	7	.639	.621	.743	.732	.802	.792
	8	.616	.604	.730	.717	.794	.780
	9	.599	.581	.708	.693	.770	.763
7	7	.616	.604	.727	.717	.787	.780
	8	.597	.581	.714	.693	.773	.763

TABLE II

VALUES OF THE POWER OF T FOR 5% SIGNIFICANCE LEVEL

$(r_1 = 0, r_2 = r)$

	r = 1		r = 2		r = 4	
δσ	T	L_1	T	L_2	T	L_4
			n = 10			
0.5	.14	.14	.16	.16	.13	.13
1.0	.35	.35	.37	.37	.28	.27
1.5	.66	.63	.64	.63	.46	.43
2.0	.88	.84	.86	.82	.64	.60
3.0	1.00	.99	.99	.98	.90	.84
			n = 20			
0.5	.19	.19	.25	.25	.26	.25
1.0	.51	.48	.68	.61	.72	.64
1.5	.91	.82	.97	.91	.97	.90
2.0	1.00	.97	1.00	.99	1.00	.99
			n = 30			
0.5	.19	.19	.26	.23	.38	.32
1.0	.58	.52	.78	.69	.91	.78
1.5	.98	.86	.99	.96	.99	.98
2.0	1.00	.99	1.00	.99	1.00	1.00

TABLE III

VALUES OF THE POWER OF T, t_c AND E_r FOR 5 AND 10 PERCENT SIGNIFICANCE LEVELS

	$r_1=r_2=1$						$r_1=r_2=2$					
	5%			10%			5%			10%		
δσ	T	t_c	E_2	T	t_c	E_2	T	t_c	E_4	T	t_c	E_4
						n = 10						
0.5	.18	.16	.13	.33	.28	.23	.19	.19	.12	.34	.31	.21
1.0	.43	.33	.31	.68	.55	.50	.42	.39	.27	.62	.61	.46
1.5	.74	.61	.59	.93	.80	.80	.68	.67	.48	.88	.84	.69
2.0	.94	.79	.85	.99	.90	.96	.89	.83	.70	.98	.94	.87
3.0	.99	.95	.99	1.00	.98	1.00	.99	.98	.92	1.00	1.00	.99
						n = 20						
0.5	.29	.19	.18	.48	.32	.34	.36	.26	.22	.56	.41	.40
1.0	.78	.52	.56	.95	.70	.74	.87	.66	.66	.96	.80	.84
1.5	.99	.79	.83	1.00	.89	.96	.99	.89	.96	1.00	.96	.99
2.0	1.00	.91	.98	1.00	.96	1.00	1.00	.98	.99	1.00	.99	1.00
						n = 30						
0.5	.32	.21	.21	.55	.33	.33	.45	.30	.30	.66	.45	.45
1.0	.91	.57	.70	.99	.73	.87	.98	.73	.86	1.00	.86	.94
1.5	1.00	.88	.98	1.00	.92	1.00	1.00	.95	.99	1.00	.98	.99

	$r_1=r_2=3$						$r_1=r_2=4$					
	5%			10%			5%			10%		
δσ	T	t_c	E_6	T	t_c	E_6	T	t_c	E_8	T	t_c	E_8
						n = 20						
0.5	.33	.30	.24	.54	.46	.40	.31	.29	.22	.50	.47	.36
1.0	.81	.71	.67	.94	.85	.82	.73	.70	.59	.90	.86	.77
1.5	.98	.94	.94	.99	.98	.97	.97	.95	.88	.99	.99	.96
2.0	1.00	.99	.99	1.00	1.00	.99	1.00	.99	.98	1.00	1.00	1.00
						n = 30						
0.5	.48	.33	.27	.69	.52	.34	.47	.38	.36	.69	.55	.51
1.0	.98	.80	.76	.99	.88	.82	.97	.85	.78	.99	.93	.91
1.5	1.00	.97	.98	1.00	.99	.99	1.00	.99	.98	1.00	1.00	.99

BIBLIOGRAPHY

[1] Bartlett, M.S. (1953). *Approximate confidence intervals.* Biometrika 40, 12-14.

[2] D'Agostino, R.B. and Tietjen, G.L. (1973). *Approaches to the null distribution on* $\sqrt{b_1}$. Biometrika 60, 169-73.

[3] David, H.A., Hartley, H.O. and Pearson, E.S. (1954). *The distribution of the ratio, in a single normal sample, of range to standard deviation.* Biometrika 41, 482-93.

[4] Dixon, W.J. (1950). *Analysis of extreme values.* Ann. Math. Statist. 21, 488-506.

[5] Dixon, W.J. (1951). *Ratios involving extreme values.* Ann. Math. Statist. 22, 68-78.

[6] Dixon, W.J. (1953). *Processing data for outliers.* Biometrics 9, 74-89.

[7] Ferguson, T.S. (1961). *On the rejection of outliers.* Proceedings Of The Fourth Berkley Symposium On Mathematical Statistics.

[8] Ferguson, T.S. (1961). *Rules for rejection of outliers.* Revue Inst. De Stat. 3, 29-43.

[9] Grubbs, F.E. (1950). *Sample criteria for testing outlying observations.* Ann. Math. Statist. 21, 27-58.

[10] Grubbs, F.E. (1969). *Procedures for detecting outlying observations in samples.* Technometrics 11, 1-21.

[11] Kendall, M.G. and Stuart, A. (1961). *The advance theory of statistics, Vol. 1.* Charless Griffin and Company Limited, London.

[12] King, E.P. (1953). *On some procedures for the rejection of suspected data.* J. Amer. Statist. Ass. 48, 531-33.

[13] Pearson, E.S. and Chandra Sekar, C. (1936). *The efficiency of statistical tools and a criterion for the rejection of outlying observations.* Biometrika 28, 308-20.

[14] Pearson, E.S. and Hartley, H.O. (1954). *Biometrika tables for statisticians, 1.* Cambridge University Press.

[15] Shaprio, S.S. and Wilk, M.B. (1965). *An analysis of variance test for normality (complete samples).* Biometrika 52, 591-611.

[16] Tietjen, G.L. and Moore, R.H. (1972). *Some Grubbs type statistics for the detection of several outliers.* Technometrics 14, 583-97.

[17] Tiku, M.L. (1967). *Estimating the mean and standard deviation from a censored normal sample.* Biometrika 54, 155-65.

[18] Tiku, M.L. (1970). *Monte Carlo study of some simple estimators in censored normal samples.* Biometrika 57, 207-11.

[19] Tiku, M.L. (1973). *Testing group effects from type II censored normal samples in experimental design.* Biometrics 29, 25-33.

[20] Tiku, M.L. (1973). *Testing normality in experimental design.* Statistics Technical Report, No. 73/1, Department of Applied Mathematics, McMaster University.

[21] Tiku, M.L. (1974). *A new statistic for testing normality.* Comm. Statist. 3(3), 223-32.

[22] Tiku, M.L. (1974). *Testing normality and exponentiality in multi-sample situations.* Comm. Statist. 3(8), 777-94.

[23] Tiku, M.L. (1975). *A new statistic for testing suspected outliers in single multi-sample situations.* Communications in Statistics (to appear).