4-2013

A Computational Method for Estimating and Finding the HConfidence Interval of the Ratio Scale Parameters in the Two-Sample Problem

Mona Abdullah Alduailij
Western Michigan University, malduailej@yahoo.com

A COMPUTATIONAL METHOD FOR ESTIMATING AND FINDING THE HCONFIDENCE INTERVAL OF THE RATIO SCALE PARAMETERS IN THE TWO-SAMPLE PROBLEM

by

Mona Abdullah Alduailij

A dissertation submitted to the The Graduate College
in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy
Department of Statistics
Western Michigan University
April 2013

Doctoral Committee:
Gerald L. Sievers, Ph.D., Chair
Jung Chao Wang, Ph.D.
Jeff Terpstra, Ph.D.
Ala Al-Fuqaha, Ph.D.

A COMPUTATIONAL METHOD FOR ESTIMATING AND FINDING THE HCONFIDENCE INTERVAL OF THE RATIO SCALE PARAMETERS IN THE TWO-SAMPLE PROBLEM

Mona Abdullah Alduailij, Ph.D.

Western Michigan University, 2013

Testing equality of variances between two samples is applied in various fields. However, in the absence of non-normal assumptions, equality of variance tests would not yield robust results. In real life situation, the absence of such assumptions is even evident, which calls for more reliable tests to accommodate for the lack of these assumptions. There are abundant parametric and nonparametric methods for estimating the scale parameter; yet a distribution-free method for estimating and finding the confident interval ratio of scale parameters in the two-sample problem would be a reliable alternative. A comparison between existing parametric and non-parametric rank tests for the two-sample scale problem will be investigated which include linear rank tests and folded rank tests with different score functions, Lehmann test, jackknife test, Sukhatme test, placement tests, permutations tests and the classical Levene tests. The developed algorithm of estimation and finding the confidence interval of the scale parameters will be examined. A Monte Carlo simulation will be used to study the performance of our algorithm under symmetric and asymmetric distributions for different sample sizes. Also, the efficiency of the proposed confidence interval will be analyzed by computing the length of the interval and its probability of coverage. In general, our algorithm performed better than the available methods for estimating the ratio of the scale parameter in the two-sample problem. This work suggests the robustness of Lehmann test and Folded Klotz test for testing equality of variances. This suggestion is supported by the proposed algorithm, which asserts that the estimator and the confidence intervals of Lehmann test and Folded Klotz test are superior compared to other tests in estimating the ratio of scale parameters in the two-sample problem. Finally, real data from a cloud-based computing environment will be analyzed.

Acknowledgments

First and foremost, I must acknowledge and thank The Almighty Allah for blessing, protecting, and guiding me throughout this period. I could never have accomplished this without the faith I have in the Almighty.

My deep gratitude is extended to my supervisors Dr. Gerald L. Sievers and Dr. Jung Chao Wang for their guidance, helpful suggestions, and coordination throughout the making of this thesis.

I also wish to thank the committee members, Dr. Terpstra and Dr. Al-Fuqaha for their help and invaluable comments.

My deep acknowledgment to the Statistics Department's chair, Dr. Magdalena Niewiadomska-Bugaj, and all members of the Statistics department.

I want to thank Dr. Ala Al-Fuqaha and Mr. Ammar Kamel, Computer Science Department of Western Michigan University, for permission to use their data in the real example.

I am most thankful to Dr. Marianne Di Pierro for proofreading, and for her helpful feedback and critical remarks

I am truly grateful to my brother-in law Faisal Allaboon, my brilliant helper, and my champion throughout the writing process. Thank you for the late-night reviews, and for your help with every obstacle in the way.

I express my deepest gratitude to my beloved parents. It's impossible to thank you adequately for everything you've done. I just wanted to tell how your guidance and unconditional love made me who I am now. From a very early age you taught me the importance of family, faith and love. You always taught me to never give up and to follow my dreams. When I was sick or scared, you would stand by me and pray for me. So no matter how much I thank you, it would never be enough, I wouldn't ask for better parents and role models. Dad may your blessed soul rest in peace, and Mom, may God bless you and always protect you as you mean the world to me.

A profound gratitude to my brothers Abdulaziz and Abdulrahman, my sisters Muneerah, Reem, Maha, and lovely Mai. Thank you for all of the wonderful memories of growing up together. Your continuous support and encouragement helped me throughout several phases of my life. Your advice helped me remember that anything is possible with hard work and determination.

I would like to thank my nieces, nephews, all relatives and friends for their help and advice. Your prayers and support have played a very important role in my life.

Mona Abdullah Alduailij

Table of Contents

List of Tables

List of Figures

Chapter 1

Hypothesis Testing

In statistical analysis there are a wide range of tests that are used to test the equality of variances between two populations. In this research we will compare existing rank tests for the two-sample scale problem and propose new methods of estimation of scale parameters. The comparison will include parametric and nonparametric tests with different score functions.

1.1 Hypothesis and Assumptions

We will test the ratio of scales of two populations, where $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ are two independent samples from the two populations. The distribution functions of these populations $F_X(x) = H(\frac{x-\theta_1}{\sigma_1})$, and $G_Y(y) = H(\frac{y-\theta_2}{\sigma_2})$, where θ's and σ's are the medians and scale parameters of the populations for X and Y, respectively. In addition to the independence assumption, the observations in each sample are mutually independent. Several of the tests add more assumptions that assume known medians (location in general), or at least equal medians if the medians are not known. But, if the medians are unknown and unequal these tests can be modified by shifting the variables by subtracting the estimated medians $\left(X_1 - \hat{\theta}_1, \ldots, X_m - \hat{\theta}_1\right)$ and $\left(Y_1 - \hat{\theta}_2, \ldots, Y_n - \hat{\theta}_2\right)$ (Hollander (1999), Fligner (1979)).

We are interested in testing the following hypothesis :

$$H_0 : \sigma = \frac{\sigma_2}{\sigma_1} = 1, \; against \; H_1 : \sigma > \frac{\sigma_2}{\sigma_1} > 1 \tag{1.1}$$

In our case, we will use the modified procedures of the tests, therefore we will assume unknown and unequal populations medians.

1.2 Tests

In the following section we will present the available parametric and nonparametic tests for the two-sample scale problem. We will refer to the combined sample size $N = m + n$, where the sample size of X's is m, and the sample size of Y's is n. The tests are F-test, permutation tests, Levene tests, linear rank tests, folded rank tests, Sukhatme test, jackknife test, placement tests. Instead of including the tables of the critical values of these tests, we will use the asymptotic null distribution to approximate the tests statistics of these tests.

1.2.1 The F-Test

The likelihood ratio test is a basic method to test the hypothesis, which starts with the assumption of normality. The likelihood ratio test, used to test the hypothesis the equality of variances from two-sample problem, results into the classical F-test. The F-test can be considered as the simplest parametric test. The test statistic is as follows:

$$F = \frac{S_2^2}{S_1^2} \tag{1.2}$$

where S_1^2 and S_2^2 are the sample variances of the first and second samples, respectively. We reject H_0 that the two variance are equal when the value of F is greater than quantiles from the F-distribution with $n-1$ and $m-1$ degrees of freedom. This test has a serious problem: when the distribution departs from normality it becomes non-robust, (Box, 1953), which means that its value is not acceptable under non-normal distributions. Pearson (1931) pointed out that the assumption of normailty of the populations could not be ignored in the F-test. Box and Andersen (1955) investigated the F-test and showed that when F-statistic normalized well, it is become asymptotically distribution free.

1.2.2 Permutation (Randomization) Test

This test was first introduced by Fisher (1935). The idea of this test is that instead of comparing the test statistic with specific known distributions, we compute the test statistic from all possible permutations of the observations. With the increasing of the fast and accurate computer calculations, the permutation tests can be done easily. There are two cases, one is called the exact permutation test where we apply the method of permutation to the classical F-test. The other one is called approximate permutation test.

1.2.2.1 Exact Permutation Test

This test can be used with any test statistic, Lehmann (1975) explained the permutations test, and he applied it to the differences of two-sample means $\overline{Y} - \overline{X}$. Due to difficulties of computing the exact all permutations, we will apply the permutation method to a finite random sample on the classical F-test.

1.2.2.2 Approximate Permutation Test

This test was suggested by Box and Anderson (1955) and the test statistic is as follows:

$$B \equiv \left(1 + \left(\frac{n-1}{m-1}F\right)^{-1}\right)^{-1} \tag{1.3}$$

where F is the classical F-test, and the degrees of freedom of permutation distribution should be transferred to a beta distribution by equating the two moments of both distributions (permutation and beta) since the third and fourth moments are the same for both distributions. Thus, the null hypothesis will be rejected if the value of B is greater than a quantile value of beta distribution, with $\frac{d(n-1)}{2}$ and $\frac{d(m-1)}{2}$ degrees of freedom, where

$$d = \left[1 + \frac{1}{2}\left(\frac{N}{N-b_2}\right)(b_2 - 3)\right]^{-1}$$

and

$$b_2 = \frac{N\left[\sum_{j=1}^{n}\left(Y_j - \overline{Y}\right)^4 + \sum_{i=1}^{m}\left(X_i - \overline{X}\right)^4\right]}{\left[\sum_{j=1}^{n}\left(Y_j - \overline{Y}\right)^2 + \sum_{i=1}^{m}\left(X_i - \overline{X}\right)^2\right]^2}$$

1.2.3 Levene Test

The Levene test is one of the most famous parametric tests in comparing variances. In (1960) Levene suggested using absolute deviations of the variables from mean which means using $\widetilde{X}_i = \left|X_i - \overline{X}\right|$ instead of X_i and $\widetilde{Y}_j = \left|Y_j - \overline{Y}\right|$ instead of Y_j , and $i = 1, \ldots, m$ and $j = 1, \ldots, n$. He also suggested using $\sqrt{\widetilde{X}_i - \widetilde{Y}_j}$, $(\widetilde{X}_i - \widetilde{Y}_j)^2$, or $\ln(\widetilde{X}_i - \widetilde{Y}_j)$ instead of absolute value. Brown and Forsythe (1974) modified the Levene test to avoid non-robustness in skewed distributions by replacing mean by median in computing the absolute deviation of variables. The modified Levene test statistic is :

$$W_{50} = \frac{\overline{B} - \overline{A}}{\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)\frac{(m-1)S_A^2 + (n-1)S_B^2}{N-2}}}$$

where $\overline{A}$, $\overline{B}$, s_A^2, and s_B^2 are the sample mean and variance for $A_i = \left|X_i - \widetilde{X}\right| i = 1, \ldots, m$, $B_j = \left|Y_j - \widetilde{Y}\right| j = 1, \ldots, n$, respectively where $\widetilde{X}$ and $\widetilde{Y}$ are the respective sample medians of the first and second samples. The null hypothsis will be rejected if $W_{50} > t_{\alpha/2,m+n-2}$. They suggested that Levene (W_{50}) can be considered as a robust test under asymmetric distributions. This modification converts the Levene test to an asymptotically distribution free test (Miller (1968)) and it is robust even in long tailed distributions (O'Brien (1996)).

In (1999) Pan suggested two more modifications on Brown and Forsythe version of Levene test. The first modification uses quantile z_α as a cutoff value instead of $t_{\alpha/2}$ to increase the power of the test and improve the size of the test. The new version is called M_{50}. The second modification suggested by Pan (1999) is using the logarithms of the mean absolute deviation from medians,

$$L_{50} = \frac{\ln \overline{B} - \ln \overline{A}}{\sqrt{\frac{1}{n}(\frac{S_A^2}{\overline{A}^2}) + \frac{1}{m}(\frac{S_B^2}{\overline{B}^2})}} \tag{1.4}$$

where $\ln \overline{A}$ and $\ln \overline{B}$ are the logarithm of $\overline{A}$ and $\overline{B}$, the test will be rejected for large values of either quantiles z_α or $t_{\alpha/2}$. Pan (1999) found that these two versions (M_{50}, L_{50}) are more powerful than W_{50}, he also showed that these two versions are asymptotically distribution-free.

1.2.4 Linear Rank Tests

Hajek and Sidak (1967) introduced a class of the linear rank tests as a test for two-sample rank tests for location and scale problems. This class of test is a linear function of the ranked samples. The form of the linear rank tests for scale problem which can be considered as a nonparametric test for dispersion, can be written as follows:

$$S_\varphi = \sum_{j=1}^{n} a\left(R_j\right) \tag{1.5}$$

where $a(i)$ are scores or weights defined by $a(i) = \varphi(\frac{i}{N+1})$, $i = 1, \ldots, N$, R_j is the rank of Y_j in the combined sample of size N, and φ-functions is defined as:

$$\varphi(u, f) = \frac{-f(F^{-1}(u))}{f(F^{-1}(u))}, \qquad 0 < u < 1 \tag{1.6}$$

where $F^{-1}(u)$ is the inverse of the cumlative distribution function of f.

By using the class of the linear rank tests, we can use one general form of these tests and each test has its own form of the score function, instead of having different equations for each test. Furthermore, under the null hypothesis and for the large sample sizes, the null mean and null variance can be used for any linear rank test, which were derived by Hajek and Sidak (1967) and Gibbons (1971) as follows:

$$E(S_\varphi) = n\bar{a}$$
$$V(S_\varphi) = \frac{mn}{N(N-1)}\sum_{i=1}^{N}(a(i)-\bar{a})^2 \tag{1.7}$$

where $\bar{a}$ is the overal mean scores. Therefore, using these moments and under the null hypothesis, the test statistic S_φ is standardized (approximated to normal distribution) as follows:

$$z = \frac{S_\varphi - E(S_\varphi)}{\sqrt{V(S_\varphi)}} \tag{1.8}$$

and we can use the result of Chernov and Savage (1958) that showed the validity of using the asymptotic normal theory under the alternative hypothesis. Another feature of this class that Hajek and Sidak (1967) have shown that the class of linear rank statistics is locally most powerful rank tests.

Hettmansperger and McKean (2011) discsued in detail the features of the score function φ as following:

1. $\varphi(u)$ is a monotone function on the interval $(0,1)$.
2. It is "squared integral"; $\int_0^1 \varphi(u) = 0$, and $\int_0^1 \varphi^2(u) = 1$.

In the class of linear rank tests the assumption about the location parameter is to be known or at least equal, since this assumption is not reliable in the real life data, it was suggested to modify these tests by subtracting the sample medians from respective samples. The resulting modified tests can be considered asymptotically distribution-free (Duran (1976)). Fligner and Hettmansperg (1979) found the limiting distribution of the modified procedures for symmetric or asymmetric distributions when the location parameter is sample median.

In this section we will discuss three different tests from the linear rank tests, which are: Mood test, Ansari-Bradley test, and Klotz test. To illustrate the idea of the score function we will include three simple examples for each test.

1.2.4.1 Mood Test

This is the first nonparametric test that deals with dispersion problems for the two-sample case, as suggested by Mood (1954). The Mood test can be written in a linear rank test form as follows:

$$T_M = \sum_{j=1}^{n} a_M(R_j) \tag{1.9}$$

where $a_M(i) = \varphi_M(\frac{i}{N+1})$, and $\varphi_M(u) = (u - \frac{1}{2})^2$ is a score function and R_j is the rank of Y_j in the combined sample.

1.2.4.2 Ansari-Bradley Test

Freund and Ansari (1957), and later Ansari and Bradley (1960) have developed a test that uses the same idea as the Mood test. Unlike the Mood test, Ansari-Bradley used the absolute value score function instead of using the quadratic . We will use the form that was suggested by Ansari and Bradley (1960) and presented by Sprent (1993):

$$T_{AB} = \sum_{j=1}^{n} a_{AB}(R_j) \tag{1.10}$$

where $a_{AB}(i) = \varphi_{AB}(\frac{i}{N+1})$, and $\varphi_{AB}(u) = \left|u - \frac{1}{2}\right|$ is a score function and R_j is the rank of Y_j in the combined sample.

1.2.4.3 Klotz Test

Klotz (1962) introduced a new version of a linear rank test by combining Mood's idea of squaring the score function, and the idea of the Van der Waerden test for location that uses the inverse of the quantiles of the standard normal distribution, which are called normal scores. The test statistic can be written as follows:

$$T_K = \sum_{j=1}^{n} a_K(R_j) \tag{1.11}$$

where $a_K(i) = \varphi_K(\frac{i}{N+1})$, and $\varphi_K(u) = \left[\Phi^{-1}(u)\right]^2$ is a score function, Φ^{-1} is the inverse cumulative distribution function (cdf) of a normal distribution, and R_j is the rank of Y_j in the combined sample.

1.2.4.4 Example

Let

$X:\quad 2 \quad 3 \quad 7 \quad 6 \quad ; m = 4$

$Y:\quad 4 \quad 9 \quad 8 \quad ; n = 3$

Using these data we will compute the linear Ansari-Bradley rank test to test the varaition between X and Y. First, we have to center the observations by subtracting the sample medians for both samples as follows:

$X - \widetilde{X} = \quad -1.5 \quad -2.5 \quad 2.5 \quad 1.5$

$Y - \widetilde{Y} \quad -4 \quad 1 \quad 0$

Next, we have to combine both sample in one vector Z and delete the zero from it, therefore, we have to change the sample size, $m = 4$, $n = 2$, and the combined sample size is $N = 6$

$Z: \quad -1.5 \quad -2.5 \quad 2.5 \quad 1.5 \quad -4 \quad 1$

After that we will order the vector Z and find the rank of these orders as follows:

$ordered\, Z: \quad -4 \quad -2.5 \quad -1.5 \quad 1 \quad 1.5 \quad 2.5$

$Rank\, of\, Z: \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6$

Now, we are ready to find the score function of the Ansari-Bradley test and compute the test as follows:

$a_{AB}(i) = \left|\frac{i}{N+1} - \frac{1}{2}\right|,$

$T_{AB} = \sum_{j=1}^{2} a_{AB}(R_j) = a(1) + a(4) = 0.429$

We will approximate the test to normal distribution and find the p-value of the test:

$E(T_{AB}) = n\overline{a}_{AB} = 2 \times 0.214 = 0.429$

$V(T_{AB}) = \frac{mn}{N(N-1)} \sum_{i=1}^{6} (a_{AB}(i) - \overline{a}_{AB})^2 = 0.0218$

$z = \frac{T_{AB} - E(T_{AB})}{\sqrt{V(T_{AB})}} = \frac{0.429 - 0.429}{0.1475} = 0$

Looking at the standard normal distribution table, we found that p-value = 0.5, therefore, we can not reject H_0 that X and Y have the same variance.

1.2.5 Folded Rank Test

Fligner and Killeen (1976) proposed a new class of nonparametric statistics. They suggested ranking the absolute value of the variables of three tests: Ansari-Bradley test, Mood test and Klotz test. They found that this class has the following properties:

1. The power of this class of tests is higher than their analogy of the linear rank tests in small sample and under symmetric distribution.

2. For equal sample sizes and under symmetric or asymmetric distributions the statistics of the non-linear rank tests are consistent for testing $H_0 : \sigma_1^2 = \sigma_2^2 \quad$ vs. $H_1 : \sigma_1^2 > \sigma_2^2$, where the medians in this case are arbitrary.

3. Under H_0, these statistics are distribution-free.

Hettmansperger and McKean (2011) named these non-linear rank tests as the Folded rank-based tests. They mentioned that the folded rank test under symmetric distributions can be considered as asymptotic distribution-free and as efficiency as the linear rank tests. Furthermore, they proposed the score functions for this class as follows:

$$\psi(u) = \left(\frac{1+u}{2}\right) \tag{1.12}$$

Later, Conover et al. (1981) suggested ranking the absolute deviation from the sample medians as: $X_i^* = \left|X_i - \widetilde{X}\right|, i = 1, \ldots, m$, and $Y_j^* = \left|Y_j - \widetilde{Y}\right|, j = 1, \ldots, n$, where $\widetilde{X}$ and $\widetilde{Y}$ are the sample medians of the first and second samples. The sample medians will be used as suggested by Conover et al. (1981) instead of grand median as suggested by Fligner and Killeen.

1.2.5.1 Folded Mood Test

Conover et al. (1981) mentioned the folded Mood test as S-R (squared rank test) which was first discussed by Conover and Iman (1978), the follows “test statistic” is the result of their research:

$$T_M = \sum_{j=1}^{n} a_M(R_j) \equiv \sum_{j=1}^{n} \left(\frac{R_j}{2(N+1)} + \frac{1}{2}\right)^2 \tag{1.13}$$

where $a_M(i) = \psi_M\left(\frac{i}{N+1}\right)$ and $\psi_M(u) = \left(\frac{1+u}{2}\right)^2$ is a score function, and R_j is the rank of Y_j^* in the combined sample.

1.2.5.2 Folded Ansari-Bradley Test

Talwar and Gentle (1977) introduced the concept for T-G named as the folded Ansari-Bradley test by Conver et al. (1981), the follows “test statistic” is the result of their research:

$$T_{AB} = \sum_{j=1}^{n} a_{AB}(R_j) \equiv \sum_{j=1}^{n} \left|\frac{R_j}{2(N+1)} + \frac{1}{2}\right| \tag{1.14}$$

where $a_{AB}(i) = \psi_{AB}\left(\frac{i}{N+1}\right)$ and $\psi_{AB}(u) = \left|\frac{1+u}{2}\right|$ is a score function, and R_j is the rank of Y_j^* in the combined sample.

1.2.5.3 Folded Klotz Test

Fligner and Killeen (1976) introduced the concept for F-K named as the folded Klotz test by Conver et al. (1981), the follows "test statistic" is the result of their research:

$$T_K = \sum_{j=1}^{n} a_K\left(R_j\right) \equiv \sum_{j=1}^{n} \left[\Phi^{-1}\left(\frac{R_j}{2(N+1)} + \frac{1}{2}\right)\right]^2 \tag{1.15}$$

where $a_K(i) = \psi_K\left(\frac{i}{N+1}\right)$ and $\psi_K(u) = \left[\Phi^{-1}\left(\frac{1+u}{2}\right)\right]^2$is the score function, Φ^{-1} is the inverse cumulative distribution function (cdf) of a normal distribution, and R_j is the rank of Y_j^* in the combined sample of size N.

1.2.6 Sukhatme Test

The Sukhatme test is one of the nonparametric tests for the two-sample scale problem. Sukhatme proposed the test in (1957) as a type of U-statistic. It is defined as follows:

$$T = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \varphi(X_i, Y_j) \tag{1.16}$$

where $\varphi(X,Y) = \begin{cases} 1 \quad if \begin{cases} either & 0 < X < Y \\ or & Y < X < 0 \end{cases} \\ 0 \quad otherwise \end{cases}$

Sukhatme (1957) gave the null mean and null variance for the test to be:

$$\begin{aligned} E(T) &= \tfrac{mn}{4}, \\ V(T) &= \frac{mn(N+7)}{48} \quad . \end{aligned} \tag{1.17}$$

He also showed that Sukhatme test is asymptotically normally distributed under null and alternative hypothesis.

1.2.7 Lehmann Test

The Lehmann test was introduced by Lehmann (1951) as another type of U-statistics, specifically of the Wilcoxon-Mann-Whietney type. The test computes the $\binom{m}{2}$ differences between X's and $\binom{n}{2}$ differences between Y's. The test statistic is of the form:

$$L = \binom{m}{2}^{-1} \binom{n}{2}^{-1} \sum_{i<j,\ k<l} \varphi\left(|X_i - X_j|, |Y_k - Y_l|\right) \tag{1.18}$$

where $\varphi(u,v) = \begin{cases} 1 & if\ u < v, \\ 0 & otherwise \end{cases}$

Sukhatme (1957) suggested that this test is not asymptotically distribution free since the variance depends on the form of the distributions of X and Y. However, in an unpublished paper Gerald Sievers found a close form for the variance of this test by estimating the value of the variance from the data, which does not depend on the form of distributions of X's and Y's. Therefore, we can say that this test is nonparametric (distribution-free) and it is asymptotically normal with mean and variance as:

$$\begin{gathered} E(L) = \tfrac{1}{2}, \\ V(L) = \tfrac{4}{m}\tfrac{1}{m}\textstyle\sum_i \left(U_1\left(X_i\right)^2\right) + \tfrac{4}{n}\tfrac{1}{n}\textstyle\sum_j \left(U_2\left(Y_j\right)^2\right) \end{gathered} \tag{1.19}$$

where

$$U_1(x) = \frac{1}{mn(n-1)/2} \sum_i \sum_{j<l} I\left(|X_i - x| < |Y_j - Y_l|\right) - \hat{\theta}$$

and

$$U_2(y) = \frac{1}{mn(m-1)/2} \sum_j \sum_{i<l} I\left(|X_i - X_l| < |Y_j - y|\right) - \hat{\theta},$$

and $\hat{\theta}$ is the estimator of θ that is defined as:

$$\hat{\theta} = \frac{\sum_{i<k}\sum_{j<l} I\left(|X_i - X_k| < |Y_j - Y_l|\right)}{m(m-1)n(n-1)/4}$$

where $\theta = P\left(|X_1 - X_2| < |Y_1 - Y_2|\right)$.

1.2.8 Jackknife Test

Miller (1968) extended the idea of Quuenuille (1949) to develop what is called the the jackknife procedure. They found that the jackknife procedure can be applied in different statistical tools. One of the applications tests the scale parameter of the two-sample problem. The test statistic is defined as:

$$Q = \frac{\overline{U} - \overline{C}}{\sqrt{V_1 + V_2}} \tag{1.20}$$

where computing Q follows jackknife procedure as follows:

For the first sample $i = 1, \ldots, m$, we will compute S_i, the natural logarithm of the sample variance for observations $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_m$. Also for the $j = 1, \ldots, n$, in the second sample we will compute T_j the natural logarithm of the sample variance for observations $Y_{1,} \ldots, Y_{j-1}, Y_{j+1}, \ldots, Y_n$. Therefore, the following equations represent the first step in a three-parts sequense:

$$D_i^2 = \frac{\sum_{s\neq i}^{m} X_s^2 - \frac{\sum_{s\neq i} X_s^2}{m-1}}{m-2},$$

$$E_j^2 = \frac{\sum_{t\neq j}^{n} Y_t^2 - \frac{\sum_{t\neq j} Y_t^2}{n-1}}{n-2} \tag{1.21}$$

Then $S_i = lnD_i^2$, for $i = 1, \ldots, m$, $T_j = lnE_j^2$, for $j = 1, \ldots, n$.

Also, the statistics for the whole samples X and Y represent the second step in a three-parts sequense as follows:

$$S_0 = ln\left[\sum_{s=1}^{m} \frac{(X_s - \overline{X}_0)^2}{m-1}\right],$$

$$T_0 = ln\left[\sum_{t=1}^{n} \frac{(Y_t - \overline{Y}_0)^2}{n-1}\right] \tag{1.22}$$

where $\overline{X}_0 = \sum_{s=1}^{m} \frac{X_s}{m}$ and $\overline{Y}_0 = \sum_{t=1}^{n} \frac{Y_t}{n}$. Next, we will compute *U's* and *C's* quantities in the following equations that represent the third step in a three-parts sequense:

$$U_i = mS_0 - m(m-1)S_i,\ for \quad i = 1, \ldots, m$$

$$C_i = nT_0 - n(n-1)T_j,\ for \quad j = 1, \ldots, n. \tag{1.23}$$

So, the quantities that are used in the statistics Q, $\overline{U}$, $\overline{C}$, V_1, and V_2 are the sample means, and sample variances for U's and C's. In this case the test will be rejected either if $Q > z_\alpha$ or if $Q > t_{\alpha,N-2}$ when the sample sizes are equal and small (Hollander and Wolfe (1999)).

1.2.9 Placements Tests

The last class of distribution-free test we will consider is the placements tests. Orban and Wolfe (1982) defined the placement tests as "procedures based on the placements of the observations in the smaller sample among the ordered observations in the larger samples." If we assume that $F_m(x)$ is the empirical distribution function of X, this concept is represented in the following equation,

$$U_j = F_m(Y_j) = \frac{(\# of X's \leq Y_j)}{m};\; j = 1, \ldots, n \tag{1.24}$$

then mU_j is the placement of Y_j through X's, Orban and Wolfe (1982).

Next, we will discuss two kinds of placements tests:

1.2.9.1 First Placement Test (Klotz)

The Klotz placement test is based on the linear rank statistic of the Klotz test (1962), and it is formulated as follows:

$$S_{n,m} = \sum_{j=1}^{n} \phi_m(U_j) \tag{1.25}$$

where $\phi_m(x) = \left\{\Phi^{-1}\left[\frac{mx+1}{m+2}\right]\right\}^2$ and Φ^{-1} is the inverse cumulative distribution function (cdf) of a normal distribution Orban and Wolfe (1982). This test is assypmtotically approximated to a normal distribution with null mean and null variance which is derived by Orban and Wolfe (1982) as follows:

$$\begin{gathered} E_0(S_{n,m}) = n\overline{\phi}_m \\ V_0(S_{n,m}) = \frac{n(N+1)}{(m+1)(m+2)}\left[\sum_{i=1}^{m} \phi_m^2\,(i/m) - (m+1)n\overline{\phi}_m^2\right] \end{gathered} \tag{1.26}$$

where $\overline{\phi}_m$ is the mean of $\phi_m(x) \equiv x$.

1.2.9.2 Second Placement Test (Wolfe)

Gillespie and Wolfe (1994) proposed another version of the placement tests as follows:

$$T_1 = \sum_{j=1}^{n} (U_j - \overline{U})^2 \tag{1.27}$$

where $\overline{U}$ is the mean of U_j's. For a large m and large n, T_1 is approximated to a normal distribution with null mean and variance which is defined by Gillespie and Wolfe (1994) as:

$$\begin{aligned} E_0(T_1) &= \tfrac{(n-1)(m+1)}{12m}, \\ V_0(T_1) &= \tfrac{(n-1)(m+1)}{360nm^3}\left[2nm(n+m+2) - n(n+1) + 3m(m+1)\right] \end{aligned} \tag{1.28}$$

References

[1] J. C. Lighthall *et al.*, Nucl. Instr. and Meth. A **622** 97–106 (2010).

Appendix

Title

The quick brown fox jumps over a lazy dog.