1 Tukey's Quick Test

Paul Mathews, Mathews Malnar and Bailey, Inc.
171 Casement Avenue, Painesville, OH 44077
Phone: 440-350-0911
E-mail: paul@mmbstatistical.com

1.1 Abstract

Tukey's Quick Test is a nonparametric hypothesis test to compare two independent samples for location. It is a nonparametric analog to the two independent sample t test. The test statistic, T, is the number of slipped data points above and below those overlapped in the two ordered data sets. When $T \geq 7$ there is sufficient evidence ($\alpha = 0.05$) to indicate that two samples have different locations. Other critical values are given and examples are presented. A variation on Tukey's Quick Test, the Tukey-Neave Test is discussed.

1.2 Key Words

hypothesis test, location, central tendency, two independent samples, nonparametric, slippage, Tukey-Neave, two sample t test, quick test

1.3 Introduction

Tukey's Quick Test is a very simple nonparametric hypothesis test used to compare two independent samples for location (median or mean). Nonparametric tests do not use an estimator (like $\bar{x}$) for a parameter (like μ) to make a decision. Rather, nonparametric tests use other statistics such as counts or rank scores of data for making decisions. In the case of Tukey's Quick Test, the two independent data sets are ordered and a count of the number of points slipped (i.e. not overlapped) is used to determine if the samples have different locations. The test is important because it is so simple to use, it does not require that any special assumptions are met, and no calculations or special tables are required.

1.4 Where the Technique is Used

Tukey's Quick Test is used to compare two independent samples for location. It can be used any time that the two independent sample t test is appropriate. It can also be used when the t test cannot if the data are not normal. It is also very easy to use with simple graphical presentations which compare two samples.

1.5 Data

1. The data must come from two independent random samples.

2. The sample sizes should be at least $n = 5$.

3. The ratio of the sample sizes (bigger/smaller) should be less than 1.33.

4. The data can be quantitative but must at least be ordinal (i.e. capable of being ordered by size).

5. The data may or may not be normally distributed.

1.6 Assumptions

1. The samples are independent and random.

2. The measurement scale is quantitative or at least ordinal.

3. The two samples are measured (or ordered) on the same scale with equal accuracy and precision.

4. The distributions of the populations being sampled are the same except for a possible difference in location.

1.7 Hypotheses Tested

Tukey's quick test is used to test for a location difference between two independent samples. The hypotheses tested are:

$$\begin{aligned} H_0 &: \tilde{\mu}_1 = \tilde{\mu}_2 \\ H_A &: \tilde{\mu}_1 \neq \tilde{\mu}_2 \end{aligned}$$

1.8 Procedure

1. Collect independent random samples of approximately the same size from the two populations.

2. Order the pooled data from the smallest value to the largest value while maintaining the identity of the data values.

3. Confirm that the data sets are slipped, that is, that one sample contains the largest value and the other sample contains the smallest value. If one sample contains both the smallest and largest values then Tukey's Quick Test cannot be used.

4. Count the number of slipped (nonoverlapping) points, T, in the ordered data sets. Ties between the two samples that begin or end the overlap are not counted. There will be points slipped below and above the region of overlap.

5. If $\begin{cases} T < 7 \text{ then accept } H_0 \\ 7 \leq T < 10 \text{ then reject } H_0 \text{ at } \alpha = 0.05 \\ 10 \leq T < 13 \text{ then reject } H_0 \text{ at } \alpha = 0.01 \\ 13 \leq T \text{ then reject } H_0 \text{ at } \alpha = 0.001 \end{cases}$

Example 1 *Light bulb walls blacken over time which decreases the light output. A design change is attempted to decrease the blackening. If the original design is A and the new design is B, samples of the two designs ordered from blackest to cleanest show:*

$$\{AAAAABAAAABBBABABABBBB\}$$

Determine if the design improvement worked.
Solution: *The ordered data shows that the two data sets are slipped. The five worst bulbs are A's and the four best bulbs are B's so the test statistic is* $T = 9$*. This means that we can reject the null hypothesis at* $\alpha = 0.05$ *and conclude that the new bulbs are cleaner.*

Example 2 *Two manufacturers, A and B, claim to make transistors with identical gain. Determine if their claim is true if the gains measured on 10 transistors selected randomly are:*

$$\{44, 41, 48, 33, 39, 51, 42, 36, 48, 47\}$$

for manufacturer A and

$$\{51, 54, 46, 53, 56, 43, 47, 50, 56, 53\}$$

for manufacturer B.
Solution: *The data sets are independent and are of the same size. The pooled data, ordered from smallest to largest, are:*

$$\{33, 36, 39, 41, 42, \mathbf{43}, 44, \mathbf{46}, 47, \mathbf{47}, 48, 48, \mathbf{50}, 51, \mathbf{51}, \mathbf{53}, \mathbf{53}, \mathbf{54}, \mathbf{56}, \mathbf{56}\}$$

where the smaller font indicates manufacturer A and the larger bold font indicates manufacturer B. The two data sets are slipped since manufacturer A has the smallest gain (33) *and manufacturer B has the largest gain* (**56**)*. The data sets are overlapped from* 43 *to* 51 *so there are* 5 *slipped points at the low end* $\{33, 36, 39, 41, 42\}$ *and* 5 *slipped points at the high end* $\{\mathbf{53}, \mathbf{53}, \mathbf{54}, \mathbf{56}, \mathbf{56}\}$*. (The pair of ties* $\{51, \mathbf{51}\}$ *determine the end of the overlapped region. The pair of ties* $\{\mathbf{56}, \mathbf{56}\}$ *are not of special concern because they come from the same sample. Both of them still count as slipped points.) The test statistic is* $T = 5+5 = 10$ *which means that we can reject the null hypothesis at* $\alpha = 0.01$ *and conclude that the transistor gains do not have the same location.*
Alternatively, it is easier to interpret the data if only symbols like A and B, or + and -, are used to indicate the two sets. In this case the data are:

$$\{AAAAABABttAABttBBBBB\}$$

where the t's indicate ties between the manufacturers (e.g. $\{51, \mathbf{51}\}$*). The number of slipped points,* $T = 5 + 5 = 10$*, is very easy to count this way.*

1.9 Discussion

1.9.1 Ties

Ties between the manufacturers in the overlap region are not of any concern, but a tie or ties at the points where the slippage begins can have an effect on the decision that is made.

While it is conservative to disqualify ties from the count of slipped points, it is safe to count a tie as 1/2 of a slipped point. In the first Example above this modification of the procedure would have changed the test statistic from $T = 10$ to $T = 10.5$ but the decision and the significance level, $\alpha = 0.01$, still wouldn't change.

Now consider the following ordered data from two sets, A and B:

$$\{AAAAttABAABBBttBB\}$$

where t indicates a tie between A and B. Under the original procedure the test statistic is $T = 6$ and the decision is that there is no difference in location between A and B. However, under the modified procedure the overlap begins and ends with ties so the test statistic is now $T = 4.5 + 2.5 = 7$ and we conclude that there is a difference between A and B.

If there are more than two points tied between the data sets at the start or end of the overlap, pairs of points can be considered to cancel each other out and then any remaining points after all pairs are dealt with count as slipped points. If the ordered data:

$$\{33, 36, 39, 41, 42, \mathbf{43}, 44, \mathbf{46}, 47, \mathbf{47}, 48, 48, \mathbf{50}, 51, \mathbf{51}, \mathbf{51}, \mathbf{53}, \mathbf{53}, \mathbf{54}, \mathbf{56}, \mathbf{56}\}$$

were observed (the only difference from the Example is the extra **51**) then the ties at 51 end the overlap but the second **51** can be counted as a slipped point. This data set gives a T statistic of $T = 11$ or $T = 11.5$ if the tie is counted as 1/2 of a slipped point. If the ordered data:

$$\{33, 36, 39, 41, 42, \mathbf{43}, 44, \mathbf{46}, 47, \mathbf{47}, 48, 48, \mathbf{50}, 51, 51, \mathbf{51}, \mathbf{53}, \mathbf{53}, \mathbf{54}, \mathbf{56}, \mathbf{56}\}$$

were observed then the test statistic would be $T = 10$ since the ties $\{51, \mathbf{51}\}$ effectively cancel each other out and the remaining 51 determines the end of the overlap.

As you can see the discussion of ties can get involved, but these issues are largely just minor details and are relatively unimportant. If all you remember is that $T \geq 7$ indicates that the two samples have different locations you understand the important and necessary part of the method.

1.9.2 Graphical Applications

Tukey's Quick Test is especially well suited for use with simple graphs. Double stem and leaf plots, dotplots, and in some cases boxplots can provide opportunities to use the test.

Example 3 *Use Tukey's Quick Test to evaluate the data in the double stem and leaf plot for a difference of location. The data are* 321, 333, 347, ...,413.

			6	8						
		4	**2**	3	5	7				
1	**3**	**7**	**4**	2	4	1	2			
320	330	340	350	360	370	380	390	400	410	420
			5	7	6	4	**5**	**5**	**2**	
				4	6	5	**3**		**3**	
					2	9				

Solution: *The slipped data values are already bolded in the plot. There are* $T = 6 + 5 = 11$ *of them so there is evidence that the two data sets have different locations at* $\alpha = 0.01$.

1.9.3 Methods of Maintaining Identity

There are many methods of maintaining identity of the data when performing the Tukey Test. If the data are numerical measurement data its possible to write the numbers down in order, from smallest to largest, but on separate lines, one data set below the other. Another method is to write the data in order but to use different colored pens or a highlighter to distinguish the samples from each other. Sometimes subscripts on the ordered numbers are used, or sometimes a bar over numbers indicates one data set and a bar under numbers indicates the other data set. The number of ways to distinguish to two data sets from each other is as diverse as the number of different kinds of data.

1.9.4 The Tukey-Neave Test

The Tukey-Neave Test is Neave's variation on Tukey's Test. It provides another test for a difference between locations when the Tukey Test fails. Proceed as in the Tukey test but discard the one point that maximizes the test statistic T without that point present. The new accept/reject criteria become:

$$\text{If}\begin{cases} T < 10 \text{ then accept } H_0 \\ 10 \le T < 13 \text{ then reject } H_0 \text{ at } \alpha = 0.05 \\ 13 \le T < 16 \text{ then reject } H_0 \text{ at } \alpha = 0.01 \\ 16 \le T \text{ then reject } H_0 \text{ at } \alpha = 0.001 \end{cases}$$

It is necessary to justify the removal of the offending data point, that is, the special cause that makes that one point different from the others in the data set must be found.

Example 4 *Students from two SAT exam preparation courses were ranked on how they performed on the SAT test. Their rankings, A for students from the first course and B for those from the second course, from worst to best were*

$$\{AABAAAAABAABABBABABBABBBB\}$$

Determine if there is a difference in performance between the students from the two courses.
Solution: *The Tukey test indicates* $T = 2 + 4 = 6$ *so there is no evidence of a difference between the two courses. However, there is one student from the B group who did much worse on the test than his fellow B students. Upon investigation it was found that the student had the flu and went blind in the middle of the exam. On this basis we can discard his performance and the new rankings become:*

$$\{AAAAAAABAABABBABABBABBBB\}$$

The new test statistic is $T = 7 + 4 = 11$ *which indicates that there is a location difference between the A and B groups at the* $\alpha = 0.05$ *level according to the Tukey-Neave test.*

1.10 References

Conover, W.J., *Practical Nonparametric Statistics*, 2nd Ed., John Wiley and Sons, 1980.
Tukey, J.W., "*A Quick, Compact, Two-Sample Test to Duckworth's Specifications,*" Technometrics, Vol. 1, No. 1, February, 1959, p. 31.
Neave, H.R., "*Quick and Simple Tests Based on Extreme Observations,*" The Journal of Quality Technology, 11:1, April 1979, p.66.