

# Package ‘basicUtils’

February 7, 2013

**Type** Package

**Title** Various utility functions that package common formulas for confidence intervals, tests of means, and other basic tasks.

**Version** 1.0

**Date** 2013-02-06

**Author** Chris Bedford

**Maintainer** Who to complain to  
<my-first-name-as-it-appears-in-author-line@buildlackey.com>

**Description** Various utility functions that package common formulas for confidence intervals, tests of means, the Marascuilo test, sample size estimation, and other basic tasks.

**License** Unlimited

## R topics documented:

marascuilo . . . . .	<a href="#">1</a>
zOrTscoreForDiffBetweenTwoMeans . . . . .	<a href="#">3</a>
<b>Index</b>	<a href="#">6</a>

---

marascuilo	<i>Marascuilo Procedure</i>
------------	-----------------------------

---

### Description

Performs the Marascuilo procedure on all pairwise combinations of proportion differences from a contingency table to see which one (if any) is significant.

### Usage

```
marascuilo(dataFrame, confidence=.95)
```

**Arguments**

**dataFrame** a data.frame with named rows and columns. The names of the groups being compared are assumed to be the columns.

**confidence** confidence level (as a percentage), used to calculate the chi squared constant.

**Value**

Returns a matrix whose rows consist the name of each pair of columns in the input data frame. The columns of the matrix provide an (a) indication as to whether or not the observed difference between the members of the pair are statistically significant ('significant' column), (b) the calculated critical range for the pair ('critical.range' column), and (c) the absolute difference between the calculated test statistic for the pair and the critical range for the pair ('abs.diff' column.) (See Examples.)

**Author(s)**

Chris Bedford, Build Lackey Labs, Inc.

**References**

<http://blog.buildlackey.com/r-code-for-executing-the-marascuilo-procedure-on-all-pairwise-combinations-of-proportion-differences-from-a-contingency-table/>

**Examples**

```
##---- Should be DIRECTLY executable !! ----
##-- ==> Define data, use random,
##--or do help(data=index) for the standard data sets.

lines <- "
      GP    PR    PP
Y      128   199   126
N      88    33    66
"

con <- textConnection(lines)
tablefoo <- read.table(con, header=TRUE)
close(con)

marascuilo(tablefoo)

#
# Should return a matrix with the following content:
#
#      pair      abs.diff      critical.range      significant
# [1,] "GP | PR" "0.265166028097063" "0.0992354018215412" "Y"
# [2,] "GP | PP" "0.0636574074074074" "0.117201905174372" "N"
# [3,] "PR | PP" "0.201508620689655" "0.100947721261772" "Y"
```

---

zOrTscoreForDiffBetweenTwoMeans

*z Or t Score For Difference Between Two Means*


---

## Description

Calculates the appropriate test statistic for the difference between two means, using either the z score formula (when both sample sizes are over 30), or the t score formula with pooled variance. Welch's correction to degrees of freedom is used when equalVariances is specified as FALSE.

## Usage

```
zOrTscoreForDiffBetweenTwoMeans(
    sampleSize1,
    u1,
    stdDev1,
    sampleSize2,
    u2,
    stdDev2,
    stdDevEstimatedFromSampleFactor = -1,
    hypothesizedDelta = 0,
    equalVariances = TRUE)
```

## Arguments

sampleSize1	Number of values in first sample
u1	Mean of values in first sample
stdDev1	Sample standard deviation calculated for values in first sample
sampleSize2	Number of values in second sample
u2	Mean of values in second sample
stdDev2	Sample standard deviation calculated for values in second sample
stdDevEstimatedFromSampleFactor	Specify -1 for this argument if the standard deviation was estimated from the samples, otherwise, if the standard deviation of the populations (for both samples) is known, then specify 0.
hypothesizedDelta	The hypothesized delta between the differences between medians in the populations, defaults to 0.
equalVariances	In cases where sampleSize1 or sampleSize2 is less than 30, if this argument is given as FALSE Welch's correction to calculate degrees of freedom will be used. Otherwise the pooled variance will be used. This argument is only relevant if sampleSize1 or sampleSize2 is less than 30.

**Value**

Returns a vector in which the first element is the calculated z or t score, and the second element is the degrees of freedom, calculated with using Welch's correction if `equalVariances` was specified as `FALSE`, otherwise the second element will always be -1 (indicating it should be ignored.)

**Author(s)**

Chris Bedford, Build Lackey Labs, Inc.

**Examples**

```
library("basicUtils")

# Unequal variance example (forces use of Welch Correction to calculate degrees of freedom)
# Uses scenario detailed on this page: http://www.unm.edu/~marcusj/2Sampletex2.pdf
#
s1= c( 19.7146,22.8245,26.3348,25.4338,20.8310,19.3516,29.1662,21.5908,25.0997,18.0220,20.8439,
      28.8265,23.8161,27.0340,23.5834,18.6316,22.4471,27.8443,25.3329,26.6790,23.7872,28.4952,
      27.9284,22.2871,13.2098 )
s2= c(40.0790,18.5252,35.8091,26.5560,31.3332,39.6987,25.1476,29.6046,24.2808,23.5064,39.7922,
      21.4682,13.1078,25.3269,30.2518,39.1803,34.6926,30.9565,29.9376,23.9296,27.6245,37.2205,33.9531,
      32.0166,37.1757,29.3769,40.7894,39.6987,27.1912,27.3089,36.1267,28.7846,26.5954,19.7374,
      33.9418,30.6148,26.8967,28.4069,30.6148,33.8551)

sampleSize1=length(s1)
sampleSize2=length(s2)
u1=mean(s1)
u2=mean(s2)
stdDev1=sd(s1)
stdDev2=sd(s2)

zOrTscoreForDiffBetweenTwoMeans ( sampleSize1, u1, stdDev1, sampleSize2, u2, stdDev2, -1, 0, FALSE)

# z score example (both populations greater than 30)
#
# From scenario described here:
# http://www.cliffsnotes.com/study\_guide/Two-Sample-z-test-for-Comparing-Two-Means.topicArticleId-267532,articleId-267532,articleId-267532

sampleSize1=75
u1=28
stdDev1=14.1

sampleSize2=50
u2=33
stdDev2=9.5

zOrTscoreForDiffBetweenTwoMeans ( sampleSize1, u1, stdDev1, sampleSize2, u2, stdDev2, -1, 0, FALSE)
```



# Index

\*Topic \textasciitildekwd1

marascuilo, [1](#)

zOrTscoreForDiffBetweenTwoMeans, [3](#)

\*Topic \textasciitildekwd2

marascuilo, [1](#)

zOrTscoreForDiffBetweenTwoMeans, [3](#)

marascuilo, [1](#)

zOrTscoreForDiffBetweenTwoMeans, [3](#)