Degrees of Freedom in T-test

Yaffa T. Atkins

Blitstein Institute, a member of the Touro College and University System[[1]](#endnote-1)

A basic problem in statistics is comparing the means of two populations. We use the T-test to check if the difference between the two populations is significant. The degrees of freedom is used to know which row to look for on the T-table. It is accepted that the degrees of freedom is calculated by adding the sample size of each population and subtracting two (df = M+N-2). [[2]](#endnote-2)[[3]](#endnote-3)[[4]](#endnote-4)[[5]](#endnote-5)

This seems surprising. If one of the populations has a small size, we expect the t-ratio to have wide tails, even if the second population has a large size. If one mean is very imprecise, how can the difference between it and the other mean be precise?

I created a simulation using R to test my theory. I created six different populations of random numbers with normal distribution that were all different sample sizes. I used 2, 3, 4, 5, 10, 40, and 100. I then compared the means between every pair of groups as well as between two different samples of each size. I then looked at the empirical distribution of the differences between the means and compared the percentiles from my distribution with what the t-table said. I ran the simulation one million times. The following tables compare the excepted degrees of freedom (using m+n-2) with the actual degrees of freedom I got from my simulation.

In the tables bellow:

* the top left box specifies what percentile is used
* num observations is the sample size of the smaller population in the comparison
* column names are the sample sizes of the larger population in the comparison.
* For each sample size, the table provides four traits
  + Actual is the empirical value from the simulations
  + Actual df is a linear interpolation between the values surrounding the actual value on the t-table
  + Expected df is the sum of the two sample sizes minus 2
  + Expected is the value on the t-table associated with the expected df





I concluded from my results that when the two samples sizes are the same, the degrees of freedom is close to n+m-2. As the two sample sizes get further and further apart, I noticed that the degrees of freedom got small compared to what was expected. If the two sample sizes a different, our expected degrees of freedom is not appropriate. I found one lecture that is consistent with my conclusion, but the author does not provide a source.[[6]](#endnote-6)

Appendix: simulation code

#set a seed so these results are reproducible

set.seed(1)

#set sample sizes for my populations

groups <- c(2,3,5,10,40,100)

#create an array to store all of the t-ratios for each pair of populations

arrayGroup <- array(rep(0, 6\*6\*1000000), dim=c(6, 6, 1000000))

#map the sample sizes into arrayGroup

map <- c(0,1,2,0,3,0,0,0,0,4,rep(0,29),5,rep(0,59),6)

#create empty arrays with one million place holders

#to store the t ratios of each pair for the one million simulations

for(i in groups){

for(j in groups){

if(j > i){

assign( sprintf("arrayGroup%sVS%s", i, j), rep(0,1000000))

}

}

}

#run the simulations one million times for more accuracy

for(b in 1:1000000){

for(a in groups){

#create a population with random N(0,1) numbers of each sample size

assign( sprintf("group%s", a), rnorm(a))

#created a second group for comparing populations of like sizes so that the mean

#and standard deviation of both groups will not be the same

assign( sprintf("groupB%s", a), rnorm(a))

}

#loop through the groups

for(i in groups){

for(j in groups){

#we only do when j >= i because below the diagonal is the same as above the diagonal

if(j >= i){

#x is set to the population of the first sample we are comparing

x <- eval(parse(text = sprintf("group%s", i)))

#y is set to the population of the second sample we are comparing

#if j == i, we use the the population in the second group

#so that the two populations will be different

if(j == i){

y <- eval(parse(text = sprintf("groupB%s", j)))

}else{

y <- eval(parse(text = sprintf("group%s", j)))

}

#w is set to sample size of group x

w <- i

#z is set to sample size of group y

z <- j

#calculate the t-ratio for comparing the two populations

sCombined <- (mean(x)-mean(y))

sdCombined <- sqrt(sd(x)^2/w + sd(y)^2/z)

ratio <- sCombined/sdCombined

#save the t-ratio in the arrayGroup using map to store it in its right place

arrayGroup[map[i],map[j],b] <- ratio

}

}

}

}

#for each pair of comparisons, find the empirical percentiles

#to compare them with the expected percentiles

for(i in groups){

for(j in groups){

if(j >= i){

assign( sprintf("quantile%sVS%s", i, j), quantile(arrayGroup[map[i],map[j],], probs = c(0.975, 0.99)))

x <- eval(parse(text = print(sprintf("quantile%sVS%s", i, j))))

print(x)

}

}

}

1. Yaffa T. Atkins is currently a student at Blitstein Institute, graduating in June. [↑](#endnote-ref-1)
2. WITTE, R. S., & WITTE, J. S. (1997). Statistics. Fort Worth: Harcourt, Brace College. Pg. 301 [↑](#endnote-ref-2)
3. Blumberger, S., & Clark, J. (1997). Statistics the easy way. Hauppauge, New York: Barron's. pg. 205 [↑](#endnote-ref-3)
4. <https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_ttest_a0000000116.htm> [↑](#endnote-ref-4)
5. <https://www.itl.nist.gov/div898/handbook/eda/section3/eda353.htm> [↑](#endnote-ref-5)
6. <http://www.stat.yale.edu/Courses/1997-98/101/meancomp.htm> [↑](#endnote-ref-6)