

Resampling Quiz

In the following document, I will walk describe several lines of code whose intention is to use resampling inference to address the question: Should I publish in ICES or CJFAS?

The code below writes a function to calculate our test statistic, the h index. The code takes a vector of values (citation.vec), creates another vector called sorted which is just citation.vec listed in decscending order, and creates another vector called paper.vec which is a vector whose values start at 1 and continue to the length of the original citation.vec. the vector hvalue calculates the number of values for which those in paper.vec are less than or equal to those in the sorted vector. Then, this h value is returned.

```
hindex <- function(citation.vec) {  
  sorted <- sort(citation.vec,decreasing=T)  
  paper.vec <- 1:length(citation.vec)  
  hvalue <- sum(paper.vec <= sorted)  
  return(hvalue)  
}
```

To test that the function works, run it with various vectors in place of citation.vec.

```
hindex(c(5,4,3,2,1,1,1,1))
```

```
## [1] 3
```

```
hindex(c(10,6,6,6,6,6,6,6,6,6,6))
```

```
## [1] 6
```

```
hindex(c(1,4,3,2,5,4,0,0,0,0,0))
```

```
## [1] 3
```

```
hindex(c(1012, 10, 5,4,3,1))
```

```
## [1] 4
```

Now, we will use the h index test statistic to see if ICES or CJFAS is a better journal pick. First we read in data from both. Then, we resample the recorded citations in both journals (X for CJFAS, Y for ICES) to get a distribution of h index values for each journal over a thousand iterations, and record the ones in which CJFAS has a better citation rating (in CJFAS.better). Then, to determine if CJFAS is better than ICES, we record the sum of CJFAS.better and divide it by the number of iterations to get a p value telling us the result. It would seem that CJFAS is not significantly better when it comes to citations.

```
ICES <- read.csv("ICES.csv")  
CJFAS <- read.csv("CJFAS.csv")  
  
niter <- 10000  
CJFAS.better <- vector(length=niter)  
for (i in 1:niter) {  
  X <- hindex(sample(CJFAS$Citations, 100, replace=F))  
  Y <- hindex(sample(ICES$Citations, 100, replace=F))  
  CJFAS.better[i] <- X>Y  
}
```

```
sum(CJFAS.better)/niter
```

```
## [1] 0.6509
```