Resampling

Abhishek Chaturvedi

May 17, 2017

KNIT YOUR DOCUMENT AS *HTML* AND SUBMIT IT AND THE `Rmd` file

Resampling

`x` is a random variable. We want to not only know what the `mean(x)` is but want to calculate the uncertainty of `mean(x)`. Measuring the uncertainty requires repeated measurements of `mean(x)`.

- Calculate the mean of `x`.
- Calculte the `sd( mean(x) )` using the **using 10-fold resampling**. Create your own folds, show your work.
- Calculte the `sd( mean(x) )` using the **using 10 bootstrap sample**. Create your own folds, show your work.

```
set.seed(1)
x <- runif(20,1,20)

#sd_cv # <- Your work here
sd_cv <- function(x) {
  data_split <- split(x, 1:10) # creating 10 equal chunks of the vector
  means <- c() # creating holder variable for mean of all subsets
  for(i in 1:10){
   means <- append(means, mean(unlist(data_split[-i]))) # removing one chunk and conv
erting list back to a vector to calculate mean
  }
  return(sd(means)) # returning SD on the mean values found above
}
#sd_cv(x)

#sd_boot # <- Your work here
sd_boot <- function(x) {
  data_split <- split(x, 1:10) #creating 10 equal chunks
  means <- c()
  for(i in 1:10){
    sampler <- sample(c(1:10), 10, replace=TRUE) #generating bootstrap chunks out of
10 chunks
    means <- append(means, mean(unlist(data_split[sampler]))) # calculating mean of a
ll the generated bootstraped chunks
  }
  return(sd(means))
}

#sd_boot(x)
```

Your answers here:

- sd_cv is: 0.4398215
- sd_boot is: 0.8566019