JK BS Example

Samuel Greeman

Since it is hard to manufacture examples of bootstrapping and jackknifing, we will be following closely with a jackknifing example from the math department at montana.edu and a bootstrapping example from ucla.edu, both of which can be found in references. First, we will start with the jackknife example:

Reading in the data

```
## [1] 1.544429
```

Jackknife the mean and bias correction

```
library(bootstrap)
mean_j <- jackknife(data.X, mean)
mean_j
```

```
## $jack.se
## [1] 0.2518403
##
## $jack.bias
## [1] 0
##
## $jack.values
##  [1] 1.41015 1.55665 1.57215 1.50115 1.52815 1.45765 1.58815 1.46565 1.58615
## [10] 1.58065 1.58815 1.57765 1.51465 1.60715 1.50115 1.46715 1.57065 1.59315
## [19] 1.59865 1.55260 1.61555
##
## $call
## jackknife(x = data.X, theta = mean)
```

```
adj_mean_j = mean(data.X) - mean_j$jack.bias
adj_mean_j
```

```
## [1] 1.544429
```

As you can, see our boot-strapped mean is the same as our sample mean. You can also see that there is no bias here in this data, but jackknifing did change our values. Let's see what happens when we do the jackknifing on the variance.

Jackknife the variance and bias correction

```r
var(data.X)
```

```
## [1] 1.331895
```

```r
var_j <- jackknife(data.X, var)
var_j
```

```
## $jack.se
## [1] 0.3873414
##
## $jack.bias
## [1] 0
##
## $jack.values
##  [1] 1.003420 1.398693 1.385007 1.360590 1.396137 1.235530 1.359739 1.264808
##  [9] 1.363516 1.372992 1.359739 1.377598 1.382392 1.315033 1.360590 1.269982
## [17] 1.386796 1.349521 1.337006 1.400518 1.290180
##
## $call
## jackknife(x = data.X, theta = var)
```

```r
adj_var_j = var(data.X) - var_j$jack.bias
adj_var_j
```

```
## [1] 1.331895
```

Once again, we see that our variance was unbiased. Again, our sample values were altered, but the result was not. Jackknifing sometimes does not tell you much, but that means that the estimator you are using is good because it is unbiased.

Let's move onto the bootstrapping example, which uses a data set from UCLA.

Read in the data

After we read in the data, we set our function that we will use to obtain statistics to bootstrap. We chose to use correlation as our statistic:

```r
cor_fun <- function(d, i){
    d2 <- d[i,]
    return(cor(d2$write, d2$math))
}
```

In this example we will use our number of samples as 500, and here, R calculates our estimate, bias, and standard error for this bootstrap:

Bootstrap execution

```
boot_results <- boot(data.Y, cor_fun, R = 500)
boot_results
```

```
##
## ORDINARY NONPARAMETRIC BOOTSTRAP
##
##
## Call:
## boot(data = data.Y, statistic = cor_fun, R = 500)
##
##
## Bootstrap Statistics :
##      original      bias    std. error
## t1* 0.6174493 0.001682461  0.03767169
```

As you can see, this method gives us a very precise estimate of all of our metrics.