

## Bootstrapping Hypothesis Test Test on the mean

The data set come from Michelson-Morley's speed of light experiments. This is their "first" experiment. The data are in units of km/sec minus 299000.

```
> speed
 [1]  850  740  900 1070  930  850  950  980  980  880 1000
 [2]  980  930  650  760
 [16]  810 1000 1000  960  960
```

H0: mean = 734.5

Ha: mean  $\neq$  734.5

Traditional test, which assumes the observations come from a normally distributed population, is the t-test:

$$T = (\bar{X} - 734.5) / (s / \sqrt{n})$$

Here

$$\bar{X} = \text{mean}(\text{speed}) = 909$$

$$n = \text{length}(\text{speed}) = 20$$

$$s = \text{sd}(\text{speed}) = 104.9260$$

$$t(\text{observed}) = 7.437503$$

P-value =  $\text{pt}(-7.437503, 19) + (1 - \text{pt}(7.437503, 19)) = 4.863622e-07$  which is considerably less than 0.05.

Now for the bootstrap, we first adjust the data so that it conforms to the null hypothesis:

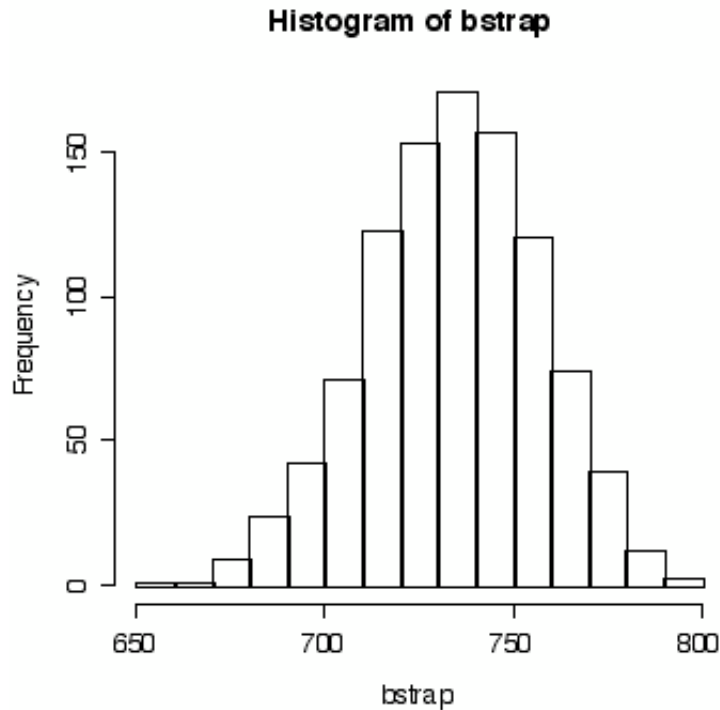
```
> newdata <- speed - mean(speed) + 734.5
> mean(newdata)
 [1] 734.5
```

Now we can proceed in either of two ways. We can either base our test on the sample mean, or we can base it on the t-statistic.

### Sample Mean:

From this "null hypothesis" distribution, we take 1000 re-samples, and each time calculate the mean. We're interested in seeing how often it's more extreme than the observed value of 909.

```
> bstrap <- c()
> for (i in 1:1000){
+   bsample <- sample(newdata, length(newdata), replace=T)
+   bstrap <- c(mean(bsample), bstrap)}
```



A picture will show us the type of averages one gets if one samples from this distribution: You can see that a value of 909 is unheard of. In fact, because it never occurred, our estimate of the p-value here would have to be 0.

#### T-test

We could also bootstrap the t-statistic:

```
myt <- function(x,mu0) (mean(x) - mu0)/(sd(x)/sqrt(length(x)))
```

The observed value of the t-test is

```
> myt(speed, 734.5)
[1] 7.437503
```

And here's the routine:

```
> bstrap <- c()
> for (i in 1:1000){
+   bsample <- sample(newdata, length(newdata), replace=T)
+   bstrap <- c(myt(bsample, 734.5),bstrap)}
> hist(bstrap)
```

Again, we can see that values as large as 7.437503 (or smaller than -7.43705) are unheard of.

Let's do one more example, since this last was a little extreme.

#### Cloud Seeding

Suppose it is "well-known" that in the area of Coral Gables Florida in which the cloud-seeding data was collected, that a "typical" amount of rain, without seeding, was 4.055 (in the units of the experiment, which was volume measured over a large region of land). The cloudseeders would claim, of course, that cloud-seeding increased this amount. Now the distribution of rain amounts is skewed towards the right, and so perhaps the median is a better measure of "typical" than the mean. If so, the hypotheses could be stated as:

H<sub>0</sub>: median = 4.055

H<sub>a</sub>: median > 4.055

After the experiment, the median of the 12 cloud-seeded days was 4.53, and we might ask if this is great enough to cast down on the null hypothesis, or is within the usual variation.

The bootstrap will come to the rescue:

First, we change the data to conform to the null hypothesis:

```
> newdata <- rain - median(rain) + 4.055
> median(newdata)
[1] 4.055
```

Next we calculate 1000 bootstrap medians:

```
> bstrap <- c()
> for (i in 1:1000){
+   bsample <- sample(newdata, length(newdata), replace=T)
+   bstrap <- c(median(bsample), bstrap)}
> hist(bstrap)
```

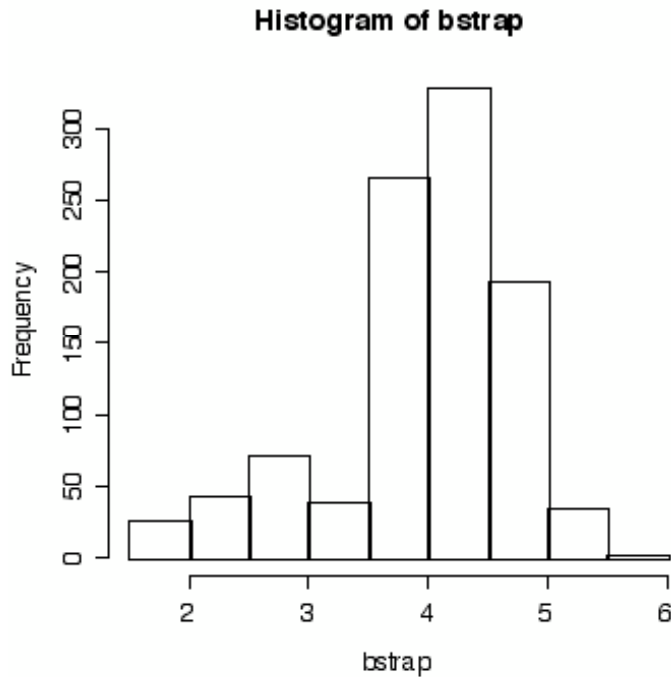
The p-value is estimated by the proportion of bootstrap medians that fell above the observed value. (We only care about those that fell above because our alternative hypothesis was that the true median was larger than the null hypothesis.)

We can find that out with these commands:

```
> sum(bstrap > 4.055)
[1] 465
```

So the proportion was  $465/1000 = .465$  or 46.5%. This indicates that this is not an unusual outcome, and so we would NOT reject the null hypothesis.

Suppose the alternative hypothesis had been:  
median  $\neq$  4.055.



Then we would have chosen to reject had the observed median been either much larger than 4.055 or much smaller. One way of dealing with this is to note that our observed median, 4.53, was .475 above the null hypothesis. So our p-value will be the proportion of bootstrapped p-values that fall more than .475 above 4.055 or .475 below 4.055.