Iliass Tiendrebeogo
ID: 30742742

Resampling Method - Seminar 3
Math 567: Winter 2016

February 18, 2016

The `jackknife` Resampling Method

The **jackknife** was a forerunner of the bootstrap, developed by Maurice Quenouille in 1949 to estimate the **bias** of an estimator $\hat{\theta}$ of a population parameter θ, and extended (and given its name) by John Tukey in 1958 to estimate the **standard error** of $\hat{\theta}$. Draper

1. **Estimation**

   To find the `jackknife` estimate of a parameter, we estimate the parameter for each subsample omitting the ith observation to estimate the previously unknown value of parameter $\bar{x}_i$

   $$\bar{x}_i = \frac{1}{n-1}\sum_{j\neq i}^{n} x_j$$

   Variance Estimation

   An estimate of the variance of an estimator can be calculated using the jackknife technique.

   $$\mathrm{Var}_{(\mathrm{jack})} = \frac{n-1}{n}\sum_{i=1}^{n}(\bar{x}_i - \bar{x}_{(.)})^2$$

   $\bar{x}_i$ is the parameter estimate based on leaving out the ith observation, and $\bar{x}_{(.)}$ is the estimator based on all of the subsamples.

   Bais Estimation

   $$\widehat{Bais}_{jack} = (n-1)(\hat{\theta}_{(.)} - \hat{\theta})$$

   first order estimate of θ is

   $$\widehat{\mathrm{Bias}}_{(\theta)} = n\hat{\theta} - (n-1)\hat{\theta}_{(.)}$$

The multiplative factor $(n-1)$ is chosen to make this work out exactly right for the case of estimating the variance. Cowles

The Relationship Between the Jackknife and the Bootstrap

The jackknife is like a bootstrap in which sampling is done without replacement instead of with, and the samples are of size $(n1)$ instead of n.

unless n is large, jackknife is less computationally intensive. But if jackknife jackknife uses fewer samples than bootstrap, then jackknife is using less information.

2. ## The jackknife estimate of bias of our dataset using R language

We are computing the ackknife estimate of bias of the median

First we install and load the `"bootrap"` package

```
install.packages("bootstrap")
library(bootstrap)
```

load the data

```
data <- read.csv('Seminar_2.csv', header = TRUE, sep = "")
```

we extract each vector from the dataset

```
col1 <- data[[1]]
col1 <- data[[2]]
```

theta function

This is the theta function which will be passed to jackknife method

```
theta <- function(x){ median(x) }
```

```
result1 <- jackknife(col1, theta)
result1$jack.bias
```

result : `> 0.1124775`

```
result2 <-jackknife(col2, theta)
result2$jack.bias
```

result : `> -0.04999`

References

Kate Cowles. Computing in statistics.

David Draper. Some notes on jackknife.