Easy 4.0 STA457 Time Series Analysis

Week Class 8 – March 16

By Yumeng Wang

1. Bootstrap

1.1 General Idea

In statistics, bootstrapping is any test or metric that relies on random sampling with replacement. Can be used to estimate bias, se, percentiles, or other properties of a statistic and compute bootstrap confidence intervals for quantities of interest

1.2 Example – Big City data

Let x be the population (in 1000's) of 49 U.S. cities in 1930. Let u be the population (in 1000's) of 49 U.S. cities in 1920. The 49 cities are a random sample taken from the 196 largest cities in 1920. We want to estimate the ratio of mean ie. $\theta = {E(x)}/{E(u)}$. We use $\bar{\theta} = {\bar{x}}/{\bar{u}}$ as the estimator.

1) Simplest but bias method

```
library(boot)
summary(bigcity)
mean(bigcity[,2])/mean(bigcity[,1])
plot(bigcity,xlab="1920 pop.", ylab="1930 pop.", cex=.6)
```

get descriptive statistics eg. mean. max. min

2) Sampling

Generate rows indices , using sample function.

```
dat<-bigcity[1:4,]; dat
set.seed(1234)
i<-sample(1:4,size=4,replace=TRUE)
i
dat[i,]
```

generate random #. different PC run out same random number → control & reproduce.

pick number randomly from 1:4. size = 4 ⇒ pick 4 times. with replacement

3) Self defined function

We want to get estimator $\bar{\theta} = {\bar{x}}/{\bar{u}}$

```
## function to compute theta.hat
ratio<-function(X) mean(X[,2])/mean(X[,1])
```

4) Bootstrap Replicates

Get random number i and resample I observations with replacement and compute estimator for each sample.

get no. of raws in dataset bigcity.

```
n<-NROW(bigcity)
t0<-ratio(bigcity)
set.seed(1234)
t.replicates<-replicate(1000, expr={
i<-sample(1:n, size=n, replace=TRUE)
boots.obs<-bigcity[i,]
ratio(boots.obs)
})
library(MASS); truehist(t.replicates); points(t0,0,pch=17,col=2)
```

return the ratio by using whole dataset

replicates 1000 times.

sample: get random number from 1:n

5) Bootstrap Standard Error

$$\widehat{var}(\hat{\theta}) = \frac{1}{R-1}\sum_{j}^{R}\left(\hat{\theta}^{(j)} - \bar{\hat{\theta}}^{*}\right)^2$$

in this case. R = 1000.

The theta star is the mean of replicates, and the thera hat is replicates getting from bootstrap

```
## Bootstrap estimate of standard error
mean(t.replicates)
se.hat<-sd(t.replicates);
se.hat
```

$\Rightarrow \bar{\hat{\theta}}^{*}$

1.3 Boot Function

<u>Basics</u>

This is more convinient way. We only need to call `boot`function.

. In the boot package in R

. syntax: `boot (data, statistic, R)`

- `Data` is the name of dataset we use
- `statistic` is the name of a function to compute the statistic of interest
- R is the number of replicates

generate set of random indices.
with placement.
from integer 1: nrow(data) and
implement the statistic function

Staticstic function

Syntax: `Function (X,i)`

. First argument to statistic is the data (matrix or dataframe)

. Second argument to statistic is the indices of the sample

. Body of function extracts the sample from the data and compute the statistic of interest. In the previous example, we are interested in the ratio.

Example

只需要这一行code. 方便.

```
out<-boot(bigcity, statistic=ratio.boot, R=999);
truehist(out$t, main="Replicate generated by boot function")
mean(out$t)
```

1.4 Bootstrap confidence interval

Method 1: use percentile

```
## the 2.5 and 97.5 percentiles determine a 95% bootstrap CI
quantile(out$t,c(0.025,0.975))
quantile(out$t,c(0.025,0.975),type=1)
## type=1 is the inverse ECDF method for computing percentiles
```

output. 2.5% 97.5% (a, b) is the CI

a b

Method 2: use confidence inverval build-in function

`boot.ci` computes five types of bootstrap confidence intervals using output from `boot: norm, perc, basic, stud, bca. Bca` produce better confidence interval

```
## boot.ci computes 5 types of bootstrap CI using output from boot
boot.ci(out,type=c("norm","perc","basic","bca"))
```

1.5 Bootstrapping Regression

Regression model:

$$y_t = \boldsymbol{X}_t\boldsymbol{\beta} + u_t, \qquad E(u_t|\boldsymbol{X}_t) = 0, \qquad u_t \sim IID(0,\sigma^2)$$

but n observations.

y_t dependent variable.

$\boldsymbol{X}_t$ matrix, contains k regressors.

$\boldsymbol{\beta}$ coefficient vector.

需要知道每个 boostrap 的方法适用于什么样的 data.

Assumptions to do different boostraping regression i.e. how to select boostrap approach.

1.5.1 Residual Boostrap

Assumptions

1. The errors u_t are independent of k regressors
2. The errors are IID, but with minimal distributional assumptions

Steps

1. use ols to get $\hat{\beta}$ & $\hat{u}_t$
2. rescale residual so that get correct variance.
   i.e. ~~centuratized.~~ centralization.
   i.e. $\ddot{u}_t = (n/n-k)^{1/2}\, \hat{u}_t$ for chosen k
3. $u_t^* \sim EDF(\ddot{u}_t)$ EDF: empirical distribution function.
   u_t^* is resampled from $\ddot{u}_t$
   $\Rightarrow$ get observation of the boostrap sample by $y_t^* = X_t\hat{\beta} + \underline{u_t^*}$

1.5.2 Parametric Bootstrap

Assumptions

u_t follows a specific distribution such as normal distribution

Steps parametric. given u_t follows a specific distribution.
usually: normal distribution.
non-parametric: don't make assumption on distribution.

steps:
1. use ols to get $\hat{\beta}$. $\hat{u}_t$
2. use $\hat{u}_t$ to get sample variance s^2.
3. observation of boostrap sample is given by $y_t^* = X_t\hat{\beta} + u_t^*$
   $u_t^* \sim NID(0, s^2)$.

bootstraping

1.5.3 Wild Bootstrap

Assumptions:

Need to handle heteroskedasticity in regression models

Model:

$y_t^* = \mathbf{X}_t\hat{\boldsymbol{\beta}} + f(\hat{u}_t)v_t^*.$

$f(\hat{u}_t)$ is function of $\hat{u}_t$ i.e. transformation of $\hat{u}_t$

V_t^* is random variable with mean 0 and variance 1

eg. $V_t^* = \pm 1$ with prob 0.5

1.6 Boostrap for dependent data

Resampling breaks up dependence within data so it is not appropriate for dependent data. Two approaches to deal with dependent data: sieve bootstrap and block bootstrap.

1.6.1 Sieve bootstrap

Suppose that the error term u_t in a regression model follow an unknown, stationary process with homoskedastic innovations. We want to use time series model e.g. AR(p) to approximate the error term.

Assumption

Assume IID innovations, thus ruling out GARCH and other forms of **heteroskedasticity**.

1.6.2 Block bootstrap

Definiton

. devide data e.g [y, X] pairs into n-b+1 blocks and each blocks have b consecutive obervations
. overlapping block and fixed size block perform better
. b increases as n increases. $b \propto n^{1/3}$

Advantage

. workes with heteroskedasticity as well as serial correlation, unlike Sieve boostrap.
. offer higher – order accuracy
. yield more reliable standard errors

Example

Consider following model

$$y_t = \boldsymbol{X}_t\boldsymbol{\beta} + \gamma y_{t-1} + u_t, \qquad u_t \sim IID(0, \sigma^2)$$

and define $\boldsymbol{Z}_t := [y_t, y_{t-1}, \boldsymbol{X}_t\,]$

So the n-b+1 fixed and overlapping blocks are

$$\boldsymbol{Z}_1, \dots, \boldsymbol{Z}_b$$

$$\boldsymbol{Z}_2, \dots, \boldsymbol{Z}_{b+1}$$

$$\dots\dots\dots$$

$$\boldsymbol{Z}_{n-b+1}, \dots, \boldsymbol{Z}_n$$

2. State space model (SSM)

It is useful tools for expressing dynamic system that involve with unobserved state variables. A state space model consists of two equations: state equation and observation equation

Observation equation: a equation that describes the relationship between observed variables y_t and state variables x_t

State equation: describe the dynamics of the state variables x_t

2.1 Univeriate state space model

observation equation: $y_t = \vec{F}^T \vec{X_t} + \varepsilon_t$

State eqn: $\vec{X_t} = \vec{G} \vec{X_{t-1}} + \vec{w_t} \Rightarrow$ VAR(1).

$y_t = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}$ univariate

$\vec{X_t}$ = state vector, $p \times 1$,

$\vec{F}$: $p \times 1$ coefficient matrix.

$\vec{G}$: $p \times p$ coefficient matrix

$\varepsilon_t \sim N(0, \sigma^2)$.

$\vec{w_t} \sim N(0, \Sigma_w)$.

2.2 SSM for AR(2)

model: $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + a_t$ $\quad a_t \sim NID(0, \sigma^2)$

AR(2).

Let $X_t = \begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix}$ $\quad w_t = \begin{bmatrix} a_t \\ 0 \end{bmatrix}$.

O: $y_t = [1 \ 0] \begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} + 0. \Rightarrow \vec{F} = [1 \quad 0]$; $\varepsilon_t = 0$.

S: $\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} a_t \\ 0 \end{pmatrix}$.

$\Rightarrow \vec{X_t} = \underbrace{\begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}}_{\vec{G}} \vec{X_{t-1}} + \vec{w_t}$. $\qquad$ not unique.

2.3 SSR for ARMA(1,1)

Model: $y_t - \phi y_{t-1} = a_t + \theta a_{t-1}$ $\quad a_t = NID(0, \sigma^2)$

$\Rightarrow y_t = \phi y_{t-1} + \theta a_{t-1} + a_t.$

$\vec{X}_t = \begin{pmatrix} y_t \\ \theta a_t \end{pmatrix} \quad \vec{w}_t = \begin{pmatrix} a_t \\ \theta a_t \end{pmatrix}$

$y_t = (1 \quad 0)\, \vec{X}_t + 0 \Rightarrow \quad \varepsilon_t = 0. \quad F = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$

$$\begin{pmatrix} y_t \\ \theta a_t \end{pmatrix} = \begin{pmatrix} \phi & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ \theta a_{t-1} \end{pmatrix} + \begin{pmatrix} a_t \\ \theta a_t \end{pmatrix}.$$

$$\vec{X}_t = \vec{G}\, \vec{X}_{t-1} + \vec{W}_t$$

2.4 Two state space models

$O: \quad y_t = \vec{F}_1' \vec{X}_t + \vec{F}_2^{T} \vec{Z}_t + \varepsilon_t$

$S_1: \quad \vec{X}_t = \vec{G}_1 \vec{X}_{t-1} + \vec{W}_t$ $\qquad$ easy!

$S_2: \quad \vec{Z}_t = \vec{G}_2 \vec{Z}_{t-1} + \vec{U}_t$

combined SSM.

$$y_t = (F_1 \quad F_2) \begin{pmatrix} X_t \\ Z_t \end{pmatrix} + \varepsilon_t.$$

$$\begin{pmatrix} X_t \\ Z_t \end{pmatrix} = \begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix} \begin{pmatrix} X_{t-1} \\ Z_{t-1} \end{pmatrix} + \begin{pmatrix} W_t \\ U_t \end{pmatrix}.$$

2.5 SSM for regression model with AR(2) error

Model: $y_t = \alpha + \beta f_t + n_t$

$n_t = \phi_1 n_{t-1} + \phi_2 n_{t-2} + a_t \qquad a_t \sim NID(0, \sigma^2).$

$\vec{X}_t = \begin{pmatrix} \alpha \\ \beta \\ n_t \\ \phi_2 n_{t-1} \end{pmatrix} \qquad y_t = [1 \quad f_t \quad 1 \quad 0]\, X_t$

$$\begin{pmatrix} \alpha \\ \beta \\ n_t \\ \phi_2 n_{t-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \phi_1 & 1 \\ 0 & 0 & \phi_2 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ n_{t-1} \\ \phi_2 n_{t-2} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ a_t \\ 0 \end{pmatrix}$$

$$X_t = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \phi_1 & 1 \\ 0 & 0 & \phi_2 & 1 \end{pmatrix}}_{\vec{G}} X_{t-1} + W_t.$$