



قدم به قدم، همراه دانشجو...

WWW.GhadamYar.Com

جامع ترین و به روزترین پرتال دانشجویی کشور (پرتال دانش)
با ارائه خدمات رایگان، تحصیلی، آموزشی، رفاهی، شغلی و...
برای دانشجویان

- (۱) راهنمای ارتقاء تحصیلی. (کاردانی به کارشناسی، کارشناسی به ارشد و ارشد به دکتری)
- (۲) ارائه سوالات کنکور مقاطع مختلف سالهای گذشته، همراه پاسخ، به صورت رایگان
- (۳) معرفی روش های مقاله و پایان نامه نویسی و ارائه پکیج های آموزشی مربوطه
- (۴) معرفی منابع و کتب مرتبط با کنکورهای تحصیلی (کاردانی تا دکتری)
- (۵) معرفی آموزشگاه ها و مراکز مشاوره تحصیلی معتبر
- (۶) ارائه جزوات و منابع رایگان مرتبط با رشته های تحصیلی
- (۷) راهنمای آزمون های حقوقی به همراه دفترچه سوالات سالهای گذشته (رایگان)
- (۸) راهنمای آزمون های نظام مهندسی به همراه دفترچه سوالات سالهای گذشته (رایگان)
- (۹) آخرین اخبار دانشجویی، در همه مقاطع، از خبرگزاری های پربازدید
- (۱۰) معرفی مراکز ورزشی، تفریحی و فروشگاه های دارای تخفیف دانشجویی
- (۱۱) معرفی همایش ها، کنفرانس ها و نمایشگاه های ویژه دانشجویی
- (۱۲) ارائه اطلاعات مربوط به بورسیه و تحصیل در خارج و معرفی شرکت های معتبر مربوطه
- (۱۳) معرفی مسائل و قوانین مربوط به سربازی، معافیت تحصیلی و امریه
- (۱۴) ارائه خدمات خاص ویژه دانشجویان خارجی
- (۱۵) معرفی انواع بیمه های دانشجویی دارای تخفیف
- (۱۶) صفحه ویژه نقل و انتقالات دانشجویی
- (۱۷) صفحه ویژه ارائه شغل های پاره وقت، اخبار استخدامی
- (۱۸) معرفی خوابگاه های دانشجویی معتبر
- (۱۹) دانلود رایگان نرم افزار و اپلیکیشن های تخصصی و...
- (۲۰) ارائه راهکارهای کارآفرینی، استارت آپ و...
- (۲۱) معرفی مراکز تاپ، ترجمه، پرینت، صحافی و ... به صورت آنلاین
- (۲۲) راهنمای خرید آنلاین ارزی و معرفی شرکت های مطرح
- (۲۳)



WWW.GhadamYar.Ir

WWW.PortaleDanesh.com

WWW.GhadamYar.Org

۰۹۱۲ ۳۰۹۰۱۰۸

باما همراه باشید...
WWW.GhadamYar.com

۰۹۱۲ ۰۹ ۰۳ ۸۰۱

پیاده سازی روش بوت استرپ در نرم افزار R

امروزه موضوع بوت استرپ به عنوان رویکرد ارتقای قابلیت یافته های آماری به شدت مورد توجه محققان مختلف است. هدف این مقاله بررسی روش شناسی آن به طور مختصر و ارائه راههای محاسبه آن در نرم افزار R است. بوت استرپ روشی ساده اما قوی از روش شبیه سازی مونت کارلو است که برای تعیین دقت آماری یا برآورد کردن توزیع از روی آماره های نمونه است. در واقع یک نوع نمونه گیری است که با تبدیل از یک نمونه به دست می آید؛ نمونه گیری درون یک نمونه. اصل برابری بوت استرپ نشان می دهد که برآوردگر نمونه گرفته شده با روش بوت استرپ، برابر است با برآورد نمونه ی اصلی. روش بوت استرپ دارای دو پیش فرض است:

- نمونه شما، نمونه معتبری از جمعیت است.
 - بوت استرپ نمونه ای را با جایگزینی از نمونه اصلی خواهد گرفت به طوری که هر نمونه گرفته شده با این روش مستقل، ولی دارای توزیع برابرند. به عبارت دیگر؛ نمونه های گرفته شده با روش بوت استرپ دارای توزیع جمعیت برابرند ولی هر نمونه، مستقل از نمونه های دیگر است.
- در اینجا مثال هایی از نمونه مسائلی که شما می توانید با روش بوت استرپ آنها را حل کنید، آورده شده:
- فرض کنید شما تعدادی داده نمونه دارید اما نمونه شما آنقدر کوچک است که شما نمی توانید درباره توزیع نمونه تان با اطمینان سخن بگویید و در نتیجه شما نمی توانید دامنه میان چارکی و یا واریانس و یا دیگر پارامترهای جمعیت تان را برآورد کنید.
 - شما دو نمونه که توزیع آن ها ناشناخته است به نام های X و Y دارید و می خواهید توزیع نسبت $Z=X/Y$ را بشناسید و بعضی از آماره های مفید Z مانند میانگین و انحراف معیار را نیز بشناسید.

- شما دو نمونه A و B دارید و می خواهید امتحان کنید که آیا آن ها از یک جمعیت یکسان هستند یا نه؟
- شما مدل رگرسیونی $Y=a+bX$ دارید و می خواهید یک فاصله اطمینان برای پارامترهای a و b بدست آورید.

بوت استرپ ناپارامتری در R: در نرم افزار R استفاده از بسته های آماری boot کار را برای انجام بازنمونه گیری بوت استرپ و روشهای شبیه سازی مشابه راحت می کند. می توانید یک آماره را (مثل میانه) یا یک بردار مثل (وزن های رگرسیون) با روش بوت استرپ به دست بیاورید. این قسمت به شما می گوید چه طور با بوت استرپ ناپارامتری اساسی شروع به کار کنید.

تابع اصلی بوت استرپ boot() با ساختار زیر است:

```
bootobject <- boot(data= , statistic= , R=, ...)
```

که در آن:

پارامتر	توصیفات
data	یک بردار vector، ماتریس matrix ویا چارچوب داده data frame
statistic	یک تابع که k آماره برای بوت استرپ شدن تولید می کند (اگر k=1 نگاه یک آماره تنها بوت استرپ می شود) این تابع باید شامل یک شاخص پارامتری باشد که تابع boot() بتواند حالتها را تکرار را انتخاب کند.
R	تعداد تکرارهای بوت استرپ
...	پارامترهای اضافی که می تواند در تابع بیابند که آماره های مورد علاقه را تولید کند.

تابع boot() آماره را R بار فرا می خواند. هر بار، این یک مجموعه از شاخص های تصادفی را تولید می کند، با جایگذاری، از اعداد صحیح (1: data) این شاخص ها در کنار تابع آماره برای انتخاب نمونه استفاده می شوند. آماره ها روی نمونه محاسبه شده و نتیجه در bootobject جمع آوری می شود. ساختار bootobject شامل موارد زیر است.

توصیفات	آرگومان
مقدار مشاهده شده از k آماره که در داده های اصلی به کار برده شدند.	$t0$
یک ماتریس $R \times k$ که هر ردیف یک تکرار بوت استرپ از k آماره می باشد.	T

می توانید به این دو آرگومان از طریق `bootobject$t0` و `t$bootobject` دست یابید.

هرگاه نمونه های بوت استرپ را تولید کردید، از `print(bootobject)` و `plot(bootobject)` برای امتحان نتیجه ها استفاده کنید. اگر نتیجه منطقی به نظر رسید، می توانید از `boot.ci()` برای رسیدن به فواصل اطمینان برای آماره تان استفاده کنید. ساختار این تابع به صورت زیر است:

`boot.ci(bootobject, conf=, type=)`

که در آن:

توصیفات	پارامتر
عملی که توسط تابع بوت استرپ برمیگردد	<code>Bootobject</code>
فاصله اطمینان مورد نظر	<code>Conf</code>
نوع فاصله اطمینان مورد نظر. مقادیر ممکن به صورت "norm"، "all"، "bca"، "perc"، "stud" و "basic" به طور پیش فرض (default: type="all")	<code>Type</code>

روش بوت استرپ برای یک آماره ی تنها؛ مثال زیر فاصله اطمینان ۹۵٪ را برای R^2 در رگرسیون خطی مایل در هر گالن (mpg) را بر وزن اتومبیل (wt) و جابجایی (disp) بررسی می کند. منبع داده ها `mtcars` می باشد. فاصله اطمینان از روش بوت استرپ بر اساس ۱۰۰۰ تکرار به دست آمد.

`Library(boot)`

`# Bootstrap 95% CI for R-Squared library(boot)`

`# function to obtain R-Squared from the data`

`rsq<-function(formula,data,indices){`

`d<-data[indices,]`

```

#allows boot to select sample

fit<-lm(formula, data=d)

return(summary(fit) $r.square)

}

#bootstrapping with 1000 replications

results<-boot(data=mtcars,statistic=rsq, R=1000,formula=mpg~wt+disp)

#viewresults

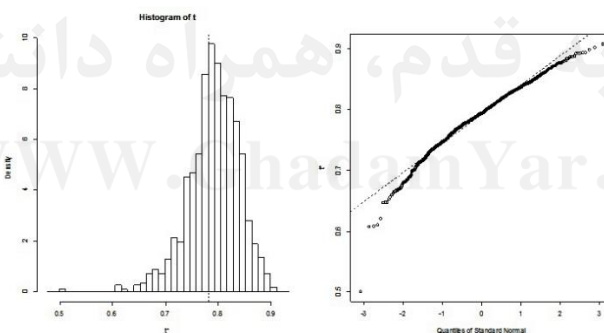
results

plot(results)

#get 95% confidence interval

boot.ci(results, type="bca")

```



بوت استرپ چندین آماره:

در مثال بالا، تابع خروجی تابع rsq یک عدد و تابع boot.ci تنها یک فاصله اطمینان به ما می دهد. تابع آماره ها می تواند به صورت یک بردار نیز باشد. در مثال بعد، ما یک فاصله اطمینان ۹۵٪ برای سه مدل

رگرسیون (جلوگیری، وزن خودرو، جابجایی) خواهیم یافت. در این حالت یک شاخص پارامتری را به `plot()` و `boot.ci()` برای نشان دادن بهتر اینکه کدام ستون در `bootobject$t` در حال تجزیه و تحلیل شدن است، اضافه می کنیم.

```
# Bootstrap 95% CI for regression coefficients library(boot)
```

```
# function to obtain regression weights
```

```
bs <- function(formula, data, indices) {
```

```
  d <- data[indices,] # allows boot to select sample
```

```
  fit <- lm(formula, data=d)
```

```
  return(coef(fit))
```

```
}
```

```
# bootstrapping with 1000 replications
```

```
results <- boot(data=mtcars, statistic=bs,
```

```
  R=1000, formula=mpg~wt+disp)
```

```
# view results
```

```
results
```

```
plot(results, index=1) # intercept
```

```
plot(results, index=2) # wt
```

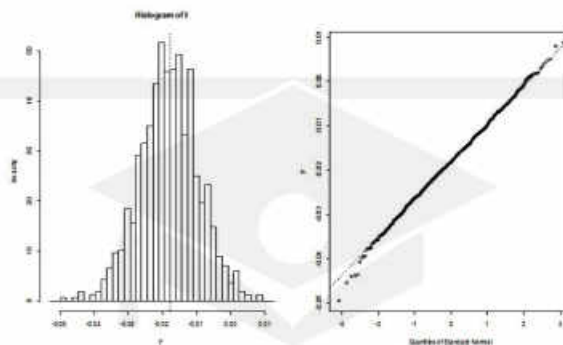
```
plot(results, index=3) # disp
```

```
# get 95% confidence intervals
```

```
boot.ci(results, type="bca", index=1) # intercept
```

```
boot.ci(results, type="bca", index=2) # wt
```

```
boot.ci(results, type="bca", index=3) # disp
```



تابع `boot()` می تواند هم به صورت پارامتری و هم ناپارامتری نمونه گیری را انجام دهد. برای بوت استرپ ناپارامتری روش باز نمونه گیری شامل موارد عادی، متعادل، ضدونقیض دار و جایگشت هاست و بازنمونه های طبقه بندی شده پشتیبانی می شوند. همچنین اهمیت وزن بازنمونه گیری می تواند مشخص گردد. تابع `boot.ci()` عامل بوت استرپ (`bootobject`) را می گیرد و ۵ نوع متفاوت از فاصله اطمینان دو طرفه را تولید می کند. شامل: اولین تخمین نرمال، فاصله بوت استرپ اصلی، فاصله بوت استرپ استیودنت، فاصله درصدی بوت استرپ، فاصله درصدی. می توانید برای کسب اطلاعات جزئیتر از `help(boot)` و `help(boot.ci)` و `help(plot.boot)` استفاده کنید.

برگرفته از

<http://amardanan.ir>