

تمرین پایان ترم
مقایسه‌ی بازه اطمینان دقیق و بوت استرپ
سارا بابایی ۹۷۱۳۰۰۵

می‌خواهیم دقت روش بازنمونه‌گیری بوت استرپ را برای محاسبه‌ی بازه اطمینان بررسی کنیم.
برای تولید نمونه‌ی تصادفی شماره دانشجویی خود را به عنوان بذر قرار دهید.

```
> set.seed(9713005)
```

یک نمونه‌ی تصادفی ۱۶ تایی از توزیع نرمال با انحراف معیار ۲ و میانگین ۵ تولید کنید.

```
> n = 16  
> mu = 5  
> sd = 2  
> x = rnorm(n, mu, sd)
```

یک بازه اطمینان ۹۵ درصد برای میانگین جامعه محاسبه کنید.

```
> u = mean(x) - z * sd / sqrt(n)  
> v = mean(x) + z * sd / sqrt(n)  
> u  
[1] 4.156244  
> v  
[1] 6.116244
```

فرض کنید انحراف معیار جامعه مجهول است. حال بازه اطمینان را دوباره محاسبه کنید.

```
> t = 2.13145  
> u1 = mean(x) - t * sqrt(var(x) / n)  
> v1 = mean(x) + t * sqrt(var(x) / n)  
> u1  
[1] 4.430781  
> v1  
[1] 5.841706
```

با دو روشی که در کتاب برای محاسبه بازه اطمینان با استفاده از روش بازنمونه‌گیری آمده دو بازه اطمینان ناپارامتری به دست آورید.

```
> b = 10000  
> boot_samples = matrix(sample(x, size=b*n, replace=TRUE),  
nrow=b, ncol=n)  
> boot_sample_mean = rep(0, b)  
> boot_sample_med = rep(0, b)  
> for (i in 1:b) {  
+   boot_sample_mean[i] = mean(boot_samples[i,])  
+   boot_sample_med[i] = median(boot_samples[i,])  
+ }  
  
> # 1. percentile bootsrtap confidence interval  
> quantile(boot_sample_mean, prob=0.025)
```

```

    2.5%
4.501381
> quantile(boot_sample_mean, prob=0.975)
    97.5%
5.769154
> quantile(boot_sample_med, prob=0.025)
    2.5%
4.362639
> quantile(boot_sample_med, prob=0.975)
    97.5%
6.136181
> # 2. basic bootsrtap confidence interval
> sort(boot_sample_mean)[b * 0.5]
[1] 5.136584
> sort(boot_sample_mean)[b * (1 - 0.5)]
[1] 5.136584

```

کدامیک از این بازه‌های اطمینان دقیق‌تر هستند؟ چرا؟
 بازه اطمینان صدکی. چون بازه‌ی بزرگ‌تری را پوشش می‌دهد.

این سوال را برای بازه اطمینان برای واریانس تکرار کنید.

```

> boot_variance = function (b, n, boot_samples) {
+   var = (b)
+   for (i in 1:b) {
+     var[i] = var(boot_samples[i,])
+   }
+   return (var)
+ }
> boot_var = boot_variance(b, n, boot_samples)
> mean(boot_var)
[1] 1.639065
> sample_var = var(x)
> sample_var
[1] 1.75274
> sort(boot_var)[b * 0.5]
[1] 1.586396
> sort(boot_var)[b * (1 - 0.5)]
[1] 1.586396

```