پروژه اول درس شبیه سازی کامپیوتری

محمدرضا اردستانی

9513004

می‌خواهیم دقت روش بازنمونه‌گیری بوت استرپ را برای محاسبه بازه اطمینان بررسی کنیم.

برای تولید نمونه تصادفی شماره دانشجویی خود را به‌عنوان بذر قرار دهید.

1) یک نمونه تصادفی ۱۶ تایی از توزیع نرمال با انحراف معیار ۲ و میانگین رقم آخر شماره دانشجویی خود تولید کنید.

2) یک بازه اطمینان ۹۵ درصد برای میانگین جامعه محاسبه کنید.

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

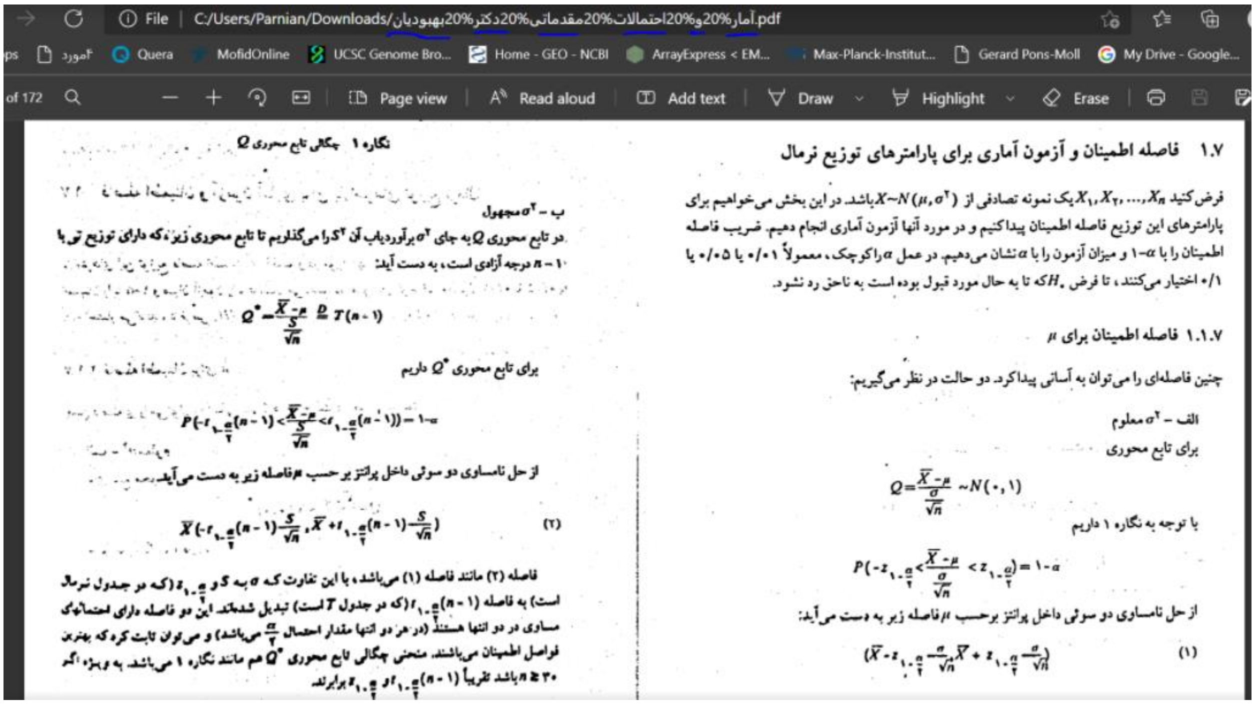
4) با دو روشی که در کتاب برای محاسبه بازۀ اطمینان با استفاده از روش بازنمونه‌گیری آمده دو بازه اطمینان ناپارامتری به دست آورید.

5) کدامیک از این بازه‌های اطمینان دقیق‌تر هستند؟ چرا؟

6) این سوال را برای بازه اطمینان برای واریانس تکرار کنید.

در کد نوشته شده خط به خط کامنت وجود دارد و دستور اجرای برنامه توضیح داده شده است. اما نکاتی که در کامنت ها امکان توضیح نبود را اینجا اضافه میکنم:

بخش دوم و سوم: فرمول های استفاده شده برای بخش 3 4 از کتاب دکتر بهبودیان و اینترنت استفاده شده است:



۱.۷ فاصله اطمینان و آزمون آماری برای پارامترهای توزیع نرمال

فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از $X\sim N(\mu,\sigma^2)$ باشد. در این بخش می‌خواهیم برای پارامترهای این توزیع فاصله اطمینان پیدا کنیم و در مورد آنها آزمون آماری انجام دهیم. ضریب فاصله اطمینان را با $1-\alpha$ و میزان آزمون را با α نشان می‌دهیم. در عمل α را کوچک، معمولاً ۰/۰۱ یا ۰/۰۵ یا ۰/۱ اختیار می‌کنند، تا فرض H_0 که تا به حال مورد قبول بوده است به ناحق رد نشود.

۱.۱.۷ فاصله اطمینان برای μ

چنین فاصله‌ای را می‌توان به آسانی پیدا کرد. دو حالت در نظر می‌گیریم:

الف – σ^2 معلوم

برای تابع محوری

$$Q=\frac{\bar X-\mu}{\frac{\sigma}{\sqrt n}}\sim N(0,1)$$

با توجه به نگاره ۱ داریم

$$P\left(-z_{1-\frac{\alpha}{2}}<\frac{\bar X-\mu}{\frac{\sigma}{\sqrt n}}<z_{1-\frac{\alpha}{2}}\right)=1-\alpha$$

از حل نامساوی دو سوئی داخل پرانتز برحسب μ فاصله زیر به دست می‌آید:

$$\left(\bar X - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt n},\ \bar X + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt n}\right) \qquad (1)$$

نگاره ۱ چگالی تابع محوری Q

ب – σ^2 مجهول

در تابع محوری Q به جای σ^2 برآوردیاب آن S^2 را می‌گذاریم تا تابع محوری زیر، که دارای توزیع تی با $n-1$ درجه آزادی است، به دست آید:

$$Q^*=\frac{\bar X-\mu}{\frac{S}{\sqrt n}} \overset{D}{=} T(n-1)$$

برای تابع محوری Q^* داریم

$$P\left(-t_{1-\frac{\alpha}{2}}(n-1)<\frac{\bar X-\mu}{\frac{S}{\sqrt n}}<t_{1-\frac{\alpha}{2}}(n-1)\right)=1-\alpha$$

از حل نامساوی دو سوئی داخل پرانتز برحسب μ فاصله زیر به دست می‌آید:

$$\left(\bar X-t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt n},\ \bar X+t_{1-\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt n}\right) \qquad (2)$$

فاصله (۲) مانند فاصله (۱) می‌باشد، با این تفاوت که σ به S و $z_{1-\frac{\alpha}{2}}$ (که در جدول نرمال است) به فاصله $t_{1-\frac{\alpha}{2}}(n-1)$ (که در جدول T است) تبدیل شده‌اند. این دو فاصله دارای احتمال‌های مساوی در دو انتها هستند (در هر دو انتها مقدار احتمال $\frac{\alpha}{2}$ می‌باشد) و می‌توان ثابت کرد که بهترین فواصل اطمینان می‌باشند. منحنی چگالی تابع محوری Q^* هم مانند نگاره ۱ می‌باشد. به ویژه اگر $n\ge 30$ باشد تقریباً $t_{1-\frac{\alpha}{2}}(n-1)$ و $z_{1-\frac{\alpha}{2}}$ برابرند.

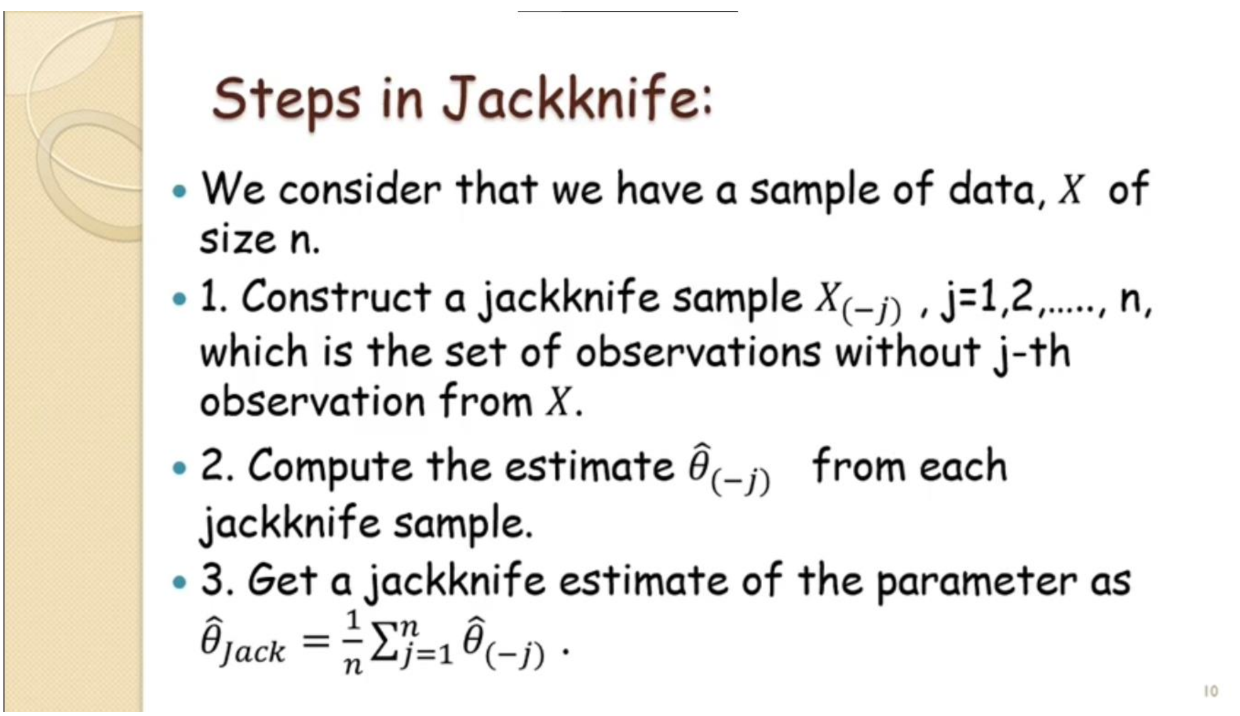
بخش چهارم: از اسلاید های Jackknife و Bootstrap کتاب و روش percentile توضیح داده شده توسط استاد استفاده شده است.

To generate a bootstrap random sample by resampling x, generate n random integers $\{i_1, \ldots, i_n\}$ uniformly distributed on $\{1, \ldots, n\}$ and select the bootstrap sample $x^* = (x_{i_1}, \ldots, x_{i_n})$.

Suppose θ is the parameter of interest (θ could be a vector), and $\hat{\theta}$ is an estimator of θ. Then the bootstrap estimate of the distribution of $\hat{\theta}$ is obtained as follows.

1. For each bootstrap replicate, indexed $b = 1, \ldots, B$:
   (a) Generate sample $x^{*(b)} = x_1^*, \ldots, x_n^*$ by sampling with replacement from the observed sample $x_1, \ldots, x_n$.
   (b) Compute the b^{th} replicate $\hat{\theta}^{(b)}$ from the b^{th} bootstrap sample.
2. The bootstrap estimate of $F_{\hat{\theta}}(\cdot)$ is the empirical distribution of the replicates $\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(B)}$.

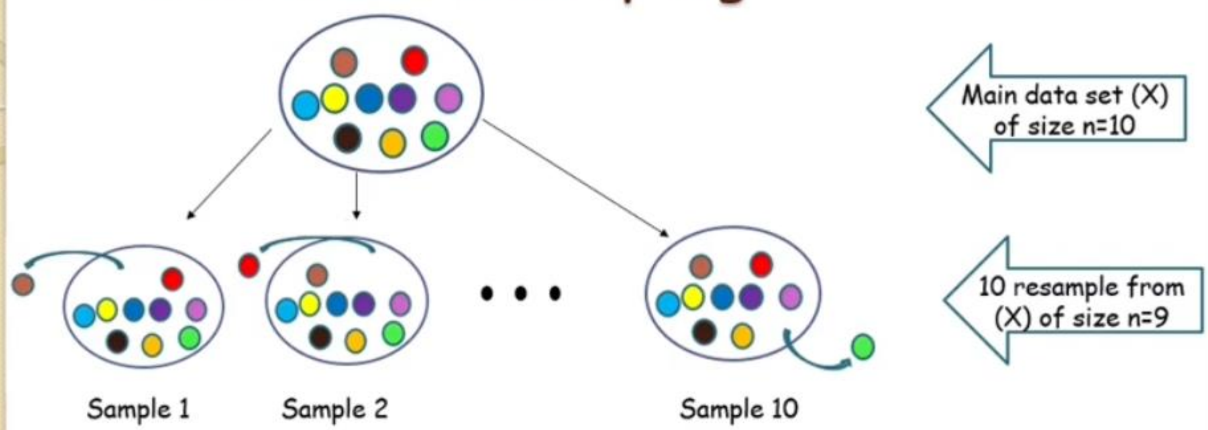
The bootstrap is applied to estimate the standard error and the bias of an estimator in the following sections. First let us see an example to illustrate the relation between the ecdf F_n and the distribution of the bootstrap replicates.



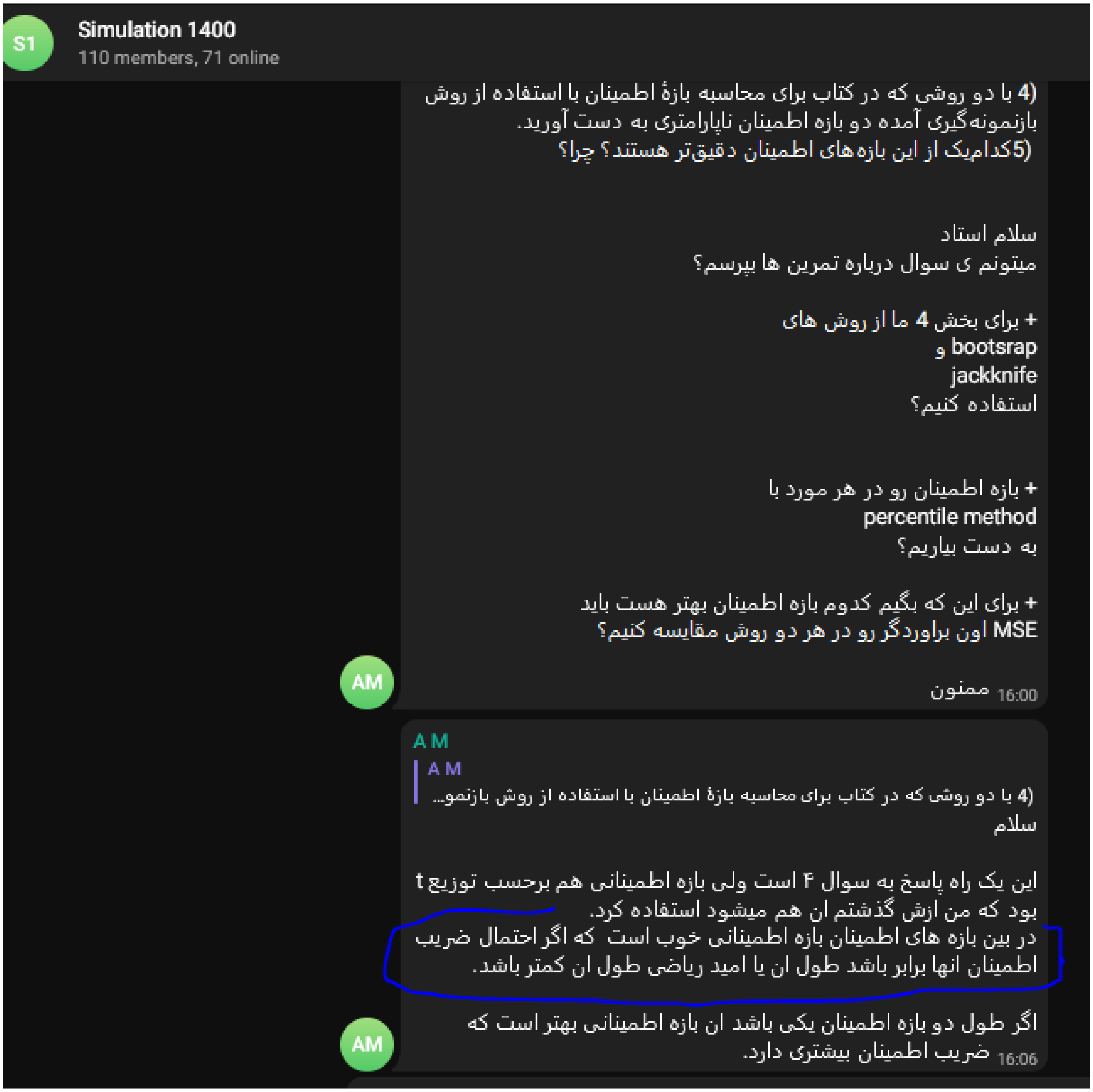
Continue…

- 4. Estimate the standard error using the formula of sample standard deviation to n estimates from Jackknife sample.
- 5. Compute other requirements, like 95% CI or bias.

Jackknife resampling



بخش پنجم: از توضیحات خود استاد در گروه استفاده شده است.


Simulation 1400

110 members, 71 online

4) با دو روشی که در کتاب برای محاسبه بازهٔ اطمینان با استفاده از روش بازنمونه‌گیری آمده دو بازه اطمینان ناپارامتری به دست آورید.
5)کدام‌یک از این بازه‌های اطمینان دقیق‌تر هستند؟ چرا؟

سلام استاد
میتونم ی سوال درباره تمرین ها بپرسم؟

+ برای بخش 4 ما از روش های
bootsrap و
jackknife
استفاده کنیم؟

+ بازه اطمینان رو در هر مورد با
percentile method
به دست بیاریم؟

+ برای این که بگیم کدوم بازه اطمینان بهتر هست باید
MSE اون براوردگر رو در هر دو روش مقایسه کنیم؟

ممنون 16:00

A M

A M

4) با دو روشی که در کتاب برای محاسبه بازهٔ اطمینان با استفاده از روش بازنمو...

سلام

این یک راه پاسخ به سوال ۴ است ولی بازه اطمینانی هم برحسب توزیع t بود که من ازش گذشتم ان هم میشود استفاده کرد.
در بین بازه های اطمینان بازه اطمینانی خوب است که اگر احتمال ضریب اطمینان انها برابر باشد طول ان یا امید ریاضی طول ان کمتر باشد.
اگر طول دو بازه اطمینان یکی باشد ان بازه اطمینانی بهتر است که ضریب اطمینان بیشتری دارد. 16:06