به نام خدا

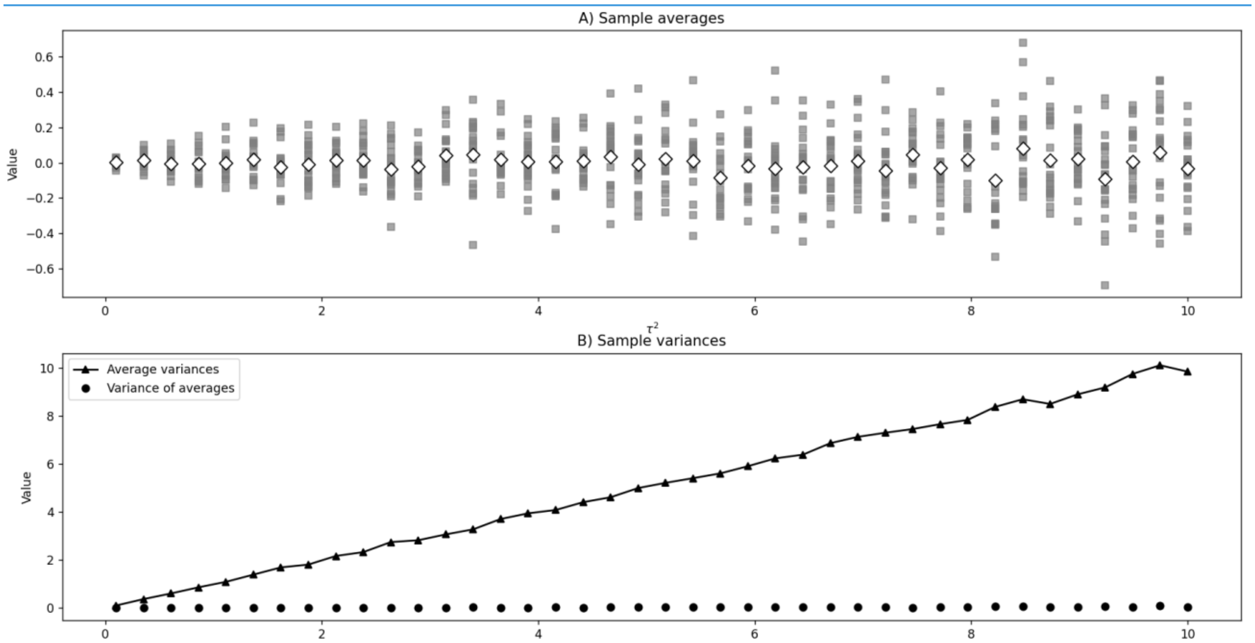
ماهان بانشی

سوال 1:

هدف سوال بررسی قانون اعداد بزرگ و محاسبه ی میانگین واریانس های داخل هر نمونه و واریانس میانگین‌های نمونه‌ها است.

در کد سوال در یک بازه ی مشخص 20 نمونه که هر کدام 200 داده دارند تولید می کنیم. سپس واریانس میانگین ها و میانگین واریانس ها را بررسی می کنیم.

نمودار های خروجی:



نتیجه: واریانس داخل نمونه‌ها با افزایش T^2 افزایش می‌یابد (کاملاً قابل انتظار است).

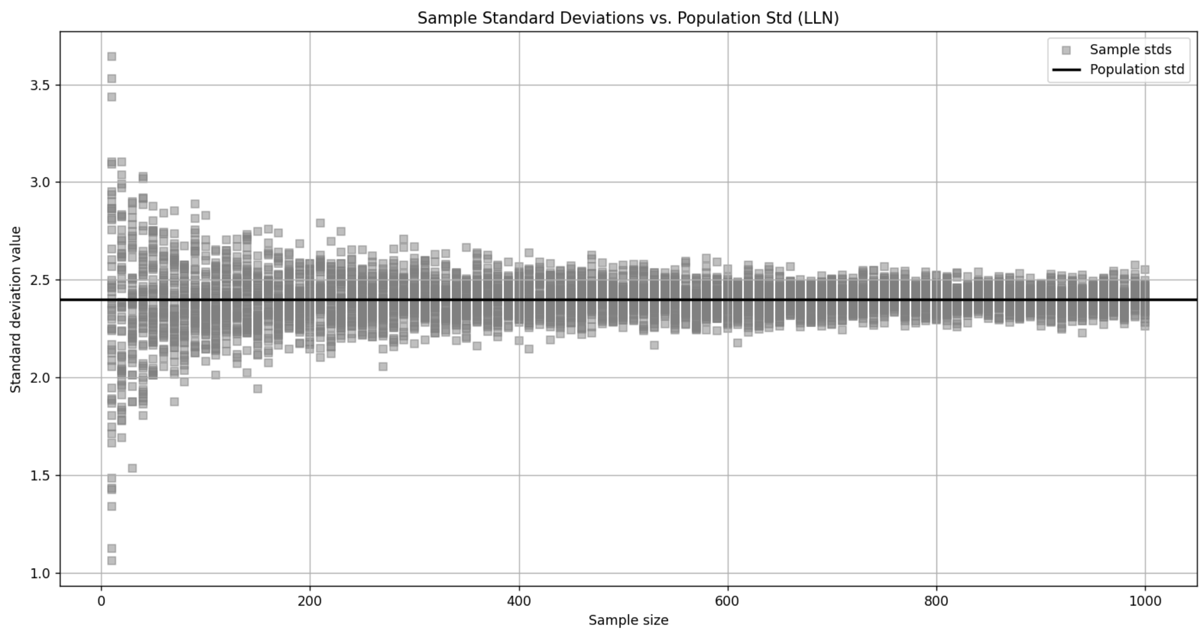
واریانس بین میانگین نمونه‌ها تقریباً ثابت باقی می‌ماند که این نشان‌دهنده‌ی پایداری میانگین‌ها حتی با افزایش نوسان داده‌ها است. این نتیجه‌ای از قانون اعداد بزرگ (LLN) است.

سوال 2:

هدف بررسی قانون اعداد بزرگ (LLN) برای انحراف معیار نمونه‌ای است (نه میانگین). با افزایش اندازه نمونه، انحراف معیار نمونه‌ای به مقدار واقعی جامعه (2.4) نزدیک می‌شود، حتی اگر LLN رسماً برای میانگین تعریف شده باشد.

در کد سوال از توزیع نرمال با میانگین 0 و انحراف معیار 2.4 نمونه گیری می کنیم و برای یک سری اندازه های مختلف انحراف معیار نمونه را حساب می کنیم. مربع های خاکستری انحراف معیار نمونه ها هستند و خط ثابت انحراف معیار جامعه است.

نمودار خروجی:



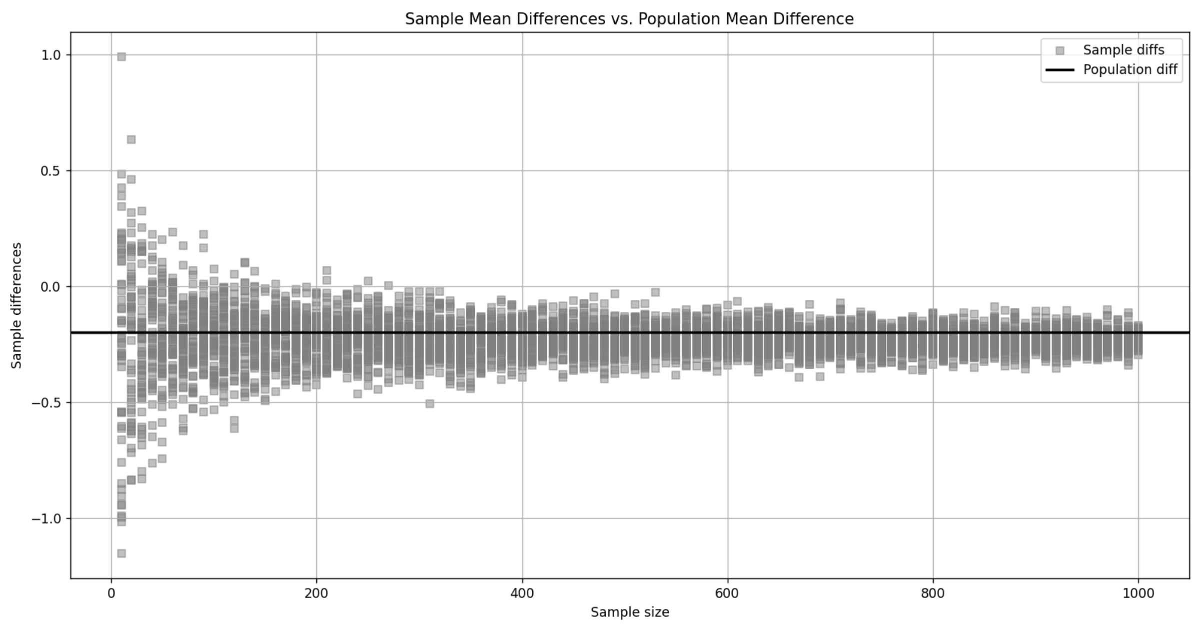
نتیجه:

هر چقدر اندازه ی نمونه ها بزرگ تر شود پراگندکی انحراف معیار کمتر می شود و مقادیر انحراف معیار به مقدار واقعی انحراف معیار جامعه نزدیک می شوند. این پدیده شبیه LLN است و نشان می‌دهد که نه فقط میانگین بلکه آمار توصیفی دیگر هم به سمت مقدار واقعی میل می‌کنند.

سوال 3:

هدف سوال مشاهده ی تغییرات در اختلاف میانگین نمونه ای بین دو جامعه ی متفاوت در اندازه های مختلف نمونه ها است.

در کد سوال دو جامعه ی نرمال با میانگین های 3 و 3.2 و انحراف معیار 1 می سازیم. سپس از هر جامعه نمونه ی تصادفی در می آوریم. سپس تفاوت میانگین نمونه ها را محاسبه می کنیم و همین کار را چندین بار انجام می دهیم و نتایج را نمایش می دهیم.



نتیجه:

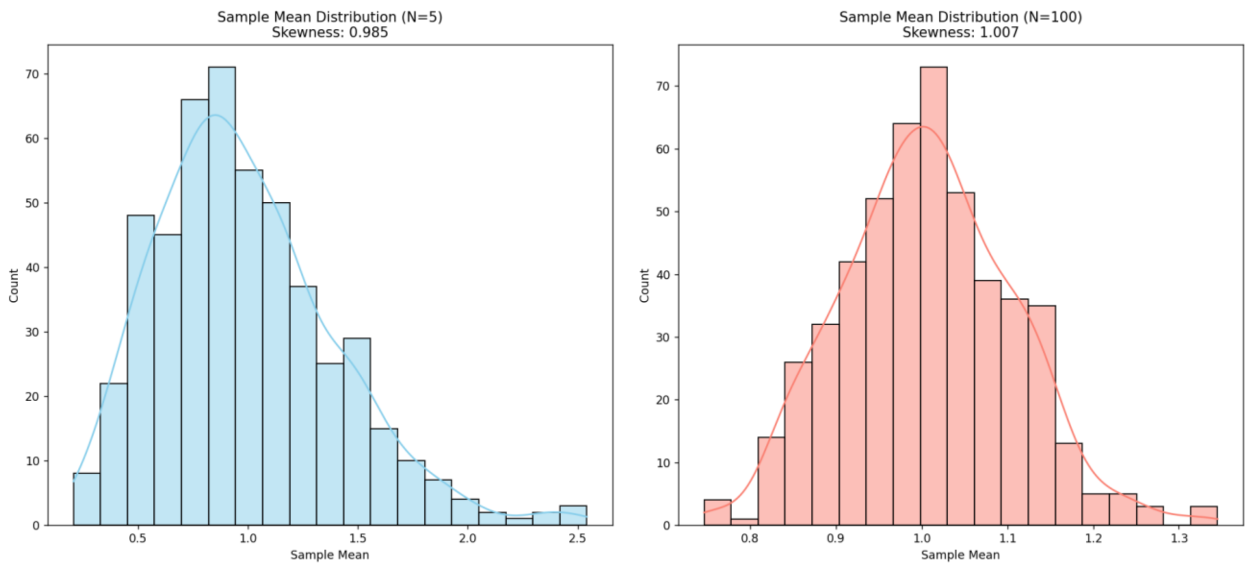
هر چه اندازه ی نمونه بیشتر شود نوسانات اختلاف میانگین نمونه ای کمتر می شود. اختلاف نمونه ها در اطراف همان مقدار واقعی که 0.2- است نزدیک تر می شود. همین هم یکی از نتیجه های مهم قانون اعداد است. یعنی با نمونه های بزرگ می توانیم با دقت بالاتری تفاوت واقعی بین دو جامعه را تخمین بزنیم.

سوال 5:

هدف کد این سوال توزیع نمونه گیری برای بررسی اهمیت اندازه ی نمونه در قضیه حد مرکزی شبیه سازی می کنیم و میانگین را برای دو اندازه نمونه مختلف بررسی می کنیم.

ما از یک توضیح نمایی با پارامتر scale =1 استفاده کرده ایم. یکبار به n 5 میدهیم و یکبار هم n را 100 می گذاریم.

نمودار خروجی:



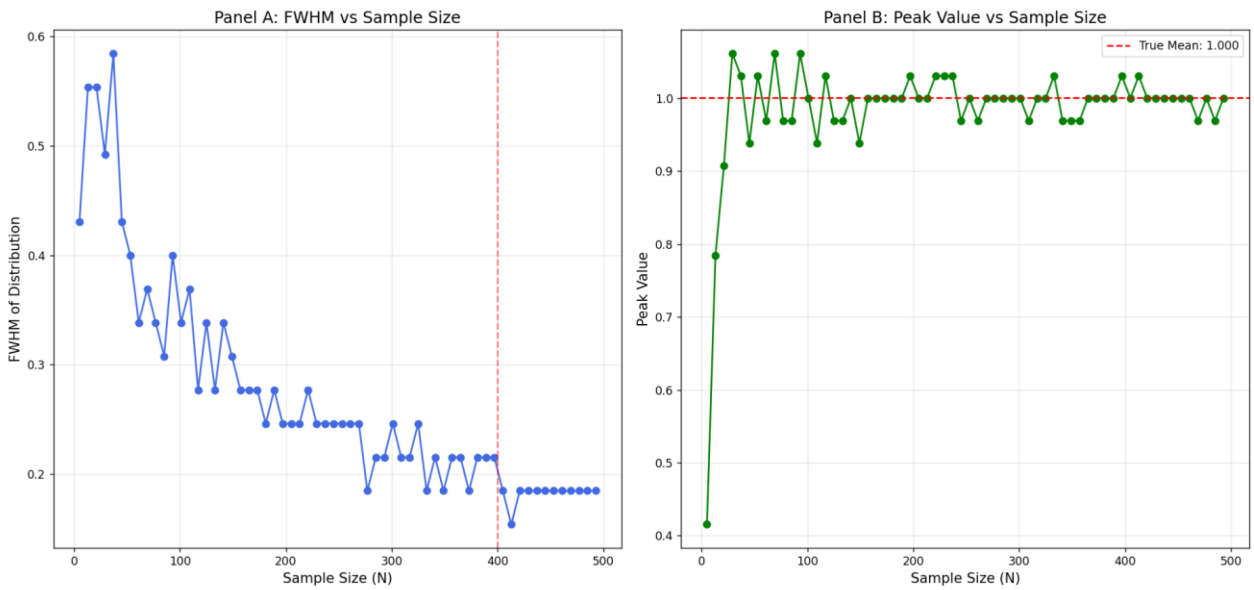
نتیجه: هرچقدر اندازه نمونه بزرگتر باشد، توزیع میانگین نمونه‌ها بیشتر به توزیع نرمال نزدیک می‌شود.

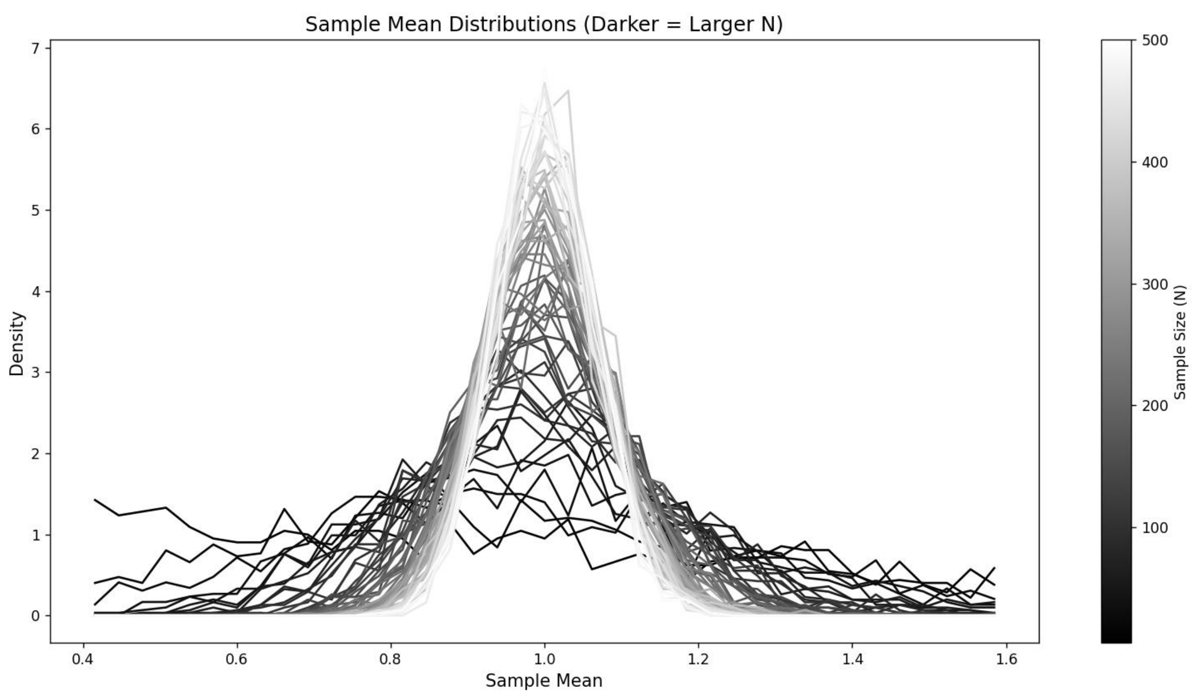
برای توزیع‌های بسیار چوله (مانند نمایی)، ممکن است به نمونه‌های بزرگتری نیاز باشد تا CLT به خوبی عمل کند. یک قاعده کلی این است که برای اکثر توزیع‌ها، N≥30 برای اعمال CLT کافی است، اما برای توزیع‌های بسیار چوله، ممکن است به نمونه‌های بزرگتری نیاز باشد.

سوال 6:

هدف این است که اندازه ی نمونه روی شکل توزیع میانگین نمونه ها از یک جمعیت مربع نرمال را بدست آوریم. در کد ابتدا جمعیت را با توزیع مربع نرمال تولید می کنیم. سپس نمونه هایی از جمعیت می گیریم و هر بار میانگینش را محاسبه می کنیم. بعد از آن یک هیستوگرام توزیع میانگین و محاسبه ی پیک می سازیم. در آخر تغییرات FWHM را بصورت گرافیکی نمایش می دهیم. و در آخر تمام توزیع ها را بر اساس شدت رنگ نسبا به اندازه ی نمونه خاکستری می کنیم.

نمودار خروجی:





نمودار اول توزیع‌های میانگین نمونه‌ها را با رنگ‌های خاکستری (تیره تر برای نمونه‌های بزرگتر) نشان می‌دهد و نمودار دوم تغییرات FWHM را بر حسب اندازه نمونه نمایش می‌دهد.

نتیجه:

1. افزایش اندازه نمونه تا N≈400 باعث می شود دقتمان بیشتر شود. (FWHM کمتر می شود)

2. پس از این نقطه، هزینه جمع‌آوری داده‌های بیشتر ممکن است توجیه پذیر نباشد.
3. میانگین توزیع نمونه‌ها حتی برای نمونه‌های نسبتاً کوچک (N > 50) به خوبی به پارامتر جمعیت نزدیک می‌شود.

سوال 7:

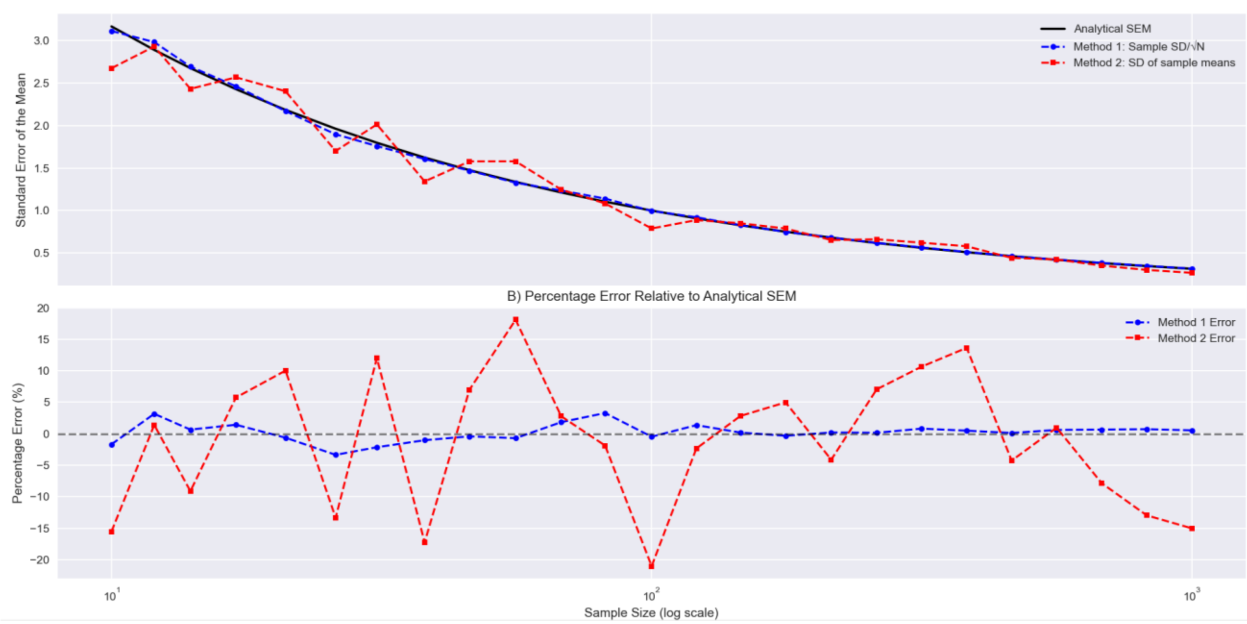
در این سوال ابتدا SEM را برای اندازه نمونه های مختلف محاسبه می کنیم. سپس با دو روش مختلف SEM را تخمین می زنیم و در آخر نتایج تئوری را با نتایج تجربی مقایسه می کنیم.

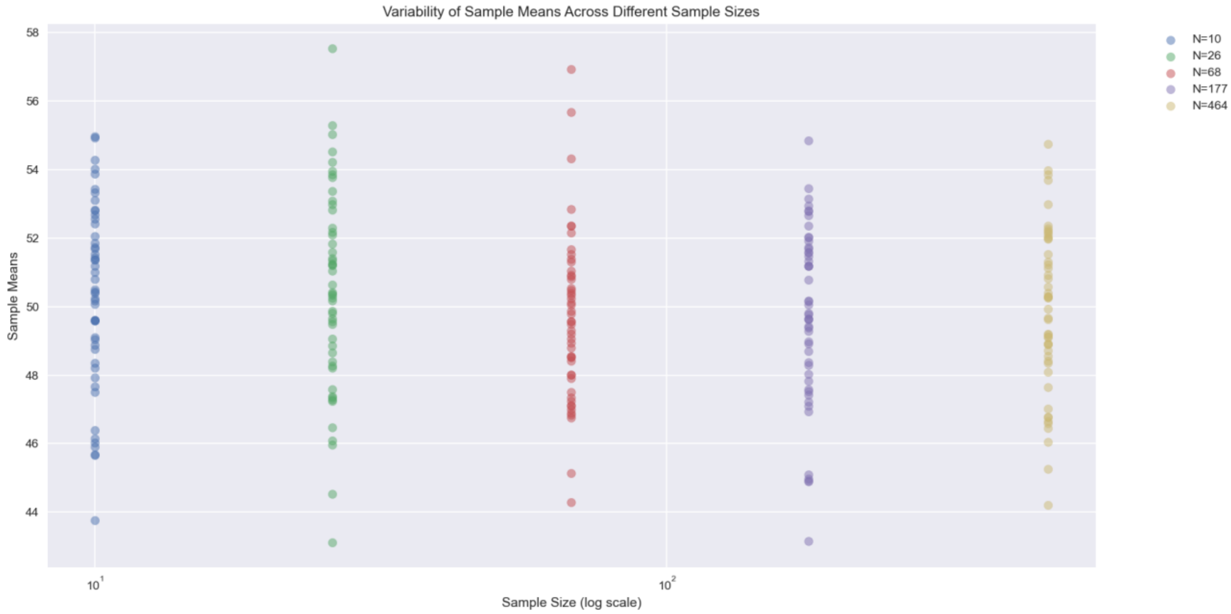
اول یک جمعیت توزیع نرمال با میانگین 50 و انحراف معیار10 ایجاد می کنیم. در ضمن SEM تحلیلی با استفاده از فرمولش محاسبه می شود. که می شود انحراف معیار تقسیم بر رادیکال n.

برای تخمین های تجربی 2 روش داریم:
1: برای هر نمونه، انحراف معیار نمونه را محاسبه و بر جذر اندازه نمونه تقسیم می‌کنیم.

2: نحراف معیار میانگین‌های نمونه‌های مختلف را محاسبه می‌کنیم.

نمودار های خروجی:





نتیجه:

1: همانطور که در خروجی مشاهده می‌شود، با افزایش اندازه نمونه، SEM کاهش می‌یابد.

2: برای نمونه‌های کوچک، اختلاف بین روش‌های تخمین و مقدار تحلیلی بیشتر است.

3: دو روش تجربی با افزایش اندازه نمونه به مقدار تحلیلی نزدیک می‌شوند.