

فهرست مطالب: بخش اول و دوم

بخش ۱: پایه‌گذاری (از احتمال به آمار)

فصل ۱: فرآیند معکوس (مدل به داده و داده به مدل)

فصل ۲: تفاوت پارامتر جامعه و آماره نمونه

فصل ۳: تصحیح جامعه متناهی و برآورد بدون سوگیری

بخش ۲: استنباط و تصمیم‌گیری (رویکرد فرکاتیستی)

فصل ۴: قضیه حد مرکزی و توزیع نمونه‌ای

فصل ۵: فاصله اطمینان و مقدمات آزمون فرض

فصل ۶: آزمون‌های یک‌دامنه و دودامنه

فصل ۷: مطالعه موردی واکسن سالک (۱۹۵۴)

فصل ۸: اخلاق آماری، P-hacking و پدیده HARKing

فهرست مطالب: بخش سوم و چهارم

بخش ۴: ابزارهای پیشرفته

- ◆ فصل ۱۴: توزیع‌های مشتق‌شده از نرمال (کای-دو و تی)
- ◆ فصل ۱۵: آزمون نیکویی برازش
- ◆ فصل ۱۶: فلوچارت عملی انتخاب آزمون

بخش ۳: تخمین پارامتر

- ◆ فصل ۹: برازش توزیع
- ◆ فصل ۱۰: روش گشتاورها
- ◆ فصل ۱۱: شبیه‌سازی مونت کارلو و بوت‌استرپ
- ◆ فصل ۱۲: برآورد بیشینه‌نمایی (MLE) مقدماتی
- ◆ فصل ۱۳: MLE عمیق و اطلاعات فشر



فهرست مطالب: بخش پنجم و ششم

بخش ۵: آمار بیزی (تغییر پارادایم)

- فصل ۱۷: مقدمه اسنتتاج بیزی
- فصل ۱۸: توزیع های مزدوج و محاسبات تحلیلی
- فصل ۱۹: خانواده نمایی و اسنتتاج بیزی
- فصل ۲۰: شیه سازی MCMC و بوت استرپ بیزی

بخش ۶: سنتز نهایی

- فصل ۲۱: رگرسیون خطی از سه دیدگاه (MLE، MAP، بیزی)

بخش اول

پایه گذاری

(از احتمال به آمار)

$$\hat{\theta}(X) \approx \theta$$

$$FPC = \sqrt{\frac{N-n}{N-1}}$$

$$E[\hat{\theta}] = \theta$$

$$s^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$$

فصل ۳:

تصحیح جامعه متناهی (FPC)
و برآورد نااریب

فصل ۲:

پارامتر جامعه و آماره نمونه

فصل ۱:

فرآیند معکوس

(مدل به داده در برابر داده به مدل)



نمونه‌گیری تصادفی و تصحیح جامعه متناهی

محدوده موضوع: درک میانگین نمونه، واریانس، و نحوه اصلاح محاسبات برای جوامع کوچک.

پیش‌نیازها: آشنایی اولیه با مفهوم میانگین (Mean) و واریانس (Variance).

اهداف یادگیری:

- ✓ درک تفاوت بین جامعه اصلی و نمونه استخراج‌شده.
- ✓ اثبات اینکه چرا میانگین نمونه، تخمین‌گری بدون سوگیری (Unbiased) است.
- ✓ محاسبه واریانس میانگین نمونه برای جوامع بزرگ (تقریب).
- ✓ اعمال «ضریب تصحیح جامعه متناهی» برای نمونه‌گیری بدون جابگذاری در جوامع کوچک.

چرا این موضوع اهمیت دارد؟

در دنیای واقعی، ما هرگز نمی‌توانیم تک‌تک افراد یک جامعه را بررسی کنیم؛ بنابراین بر اساس یک نمونه کوچک، تصمیمات بزرگ و دقیقی بگیریم. این مفاهیم پایه‌ای به ما اطمینان می‌دهند که خطای تخمین ما چقدر است و چگونه می‌توان آن را دقیق‌تر محاسبه کرد.

⚠ اشتباه رایج:

شروع به نمونه‌گیری و محاسبه بدون دانستن اینکه آیا حجم جامعه نسبت به نمونه ما بی‌نهایت محسوب می‌شود یا محدود!



جامعه در برابر نمونه (Population vs. Sample)

جامعه (N , Population)



(A) مفهوم

- جامعه (N): کل گروهی که می‌خواهیم درباره آن نتیجه‌گیری کنیم.
- نمونه (n): زیرمجموعه کوچکی از جامعه که واقعاً آن را اندازه‌گیری می‌کنیم (همیشه $n < N$).

(C) نحوه کار

3. ویژگی‌های نمونه را بسنجید تا ویژگی‌های جامعه را حدس بزنید.

2. به صورت تصادفی تعدادی را انتخاب کنید (نمونه با حجم n).

1. کل افراد یا داده‌های ممکن را تعریف کنید (جامعه با حجم N).

نمونه (n , Sample)



(B) شهود

فرض کنید می‌خواهید طعم یک دیگ بزرگ سوپ را بچشید. دیگ سوپ، «جامعه» است و یک قاشق از آن، «نمونه». نیازی نیست کل دیگ را بخورید تا بفهمید نمک آن کافی است یا خیر.

(D) مثال کاربردی

یک کارخانه در روز ۱۰۰,۰۰۰ قطعه تولید می‌کند (N). مهندس کنترل کیفیت فقط ۵۰ قطعه (n) را برای بررسی خرابی انتخاب می‌کند.

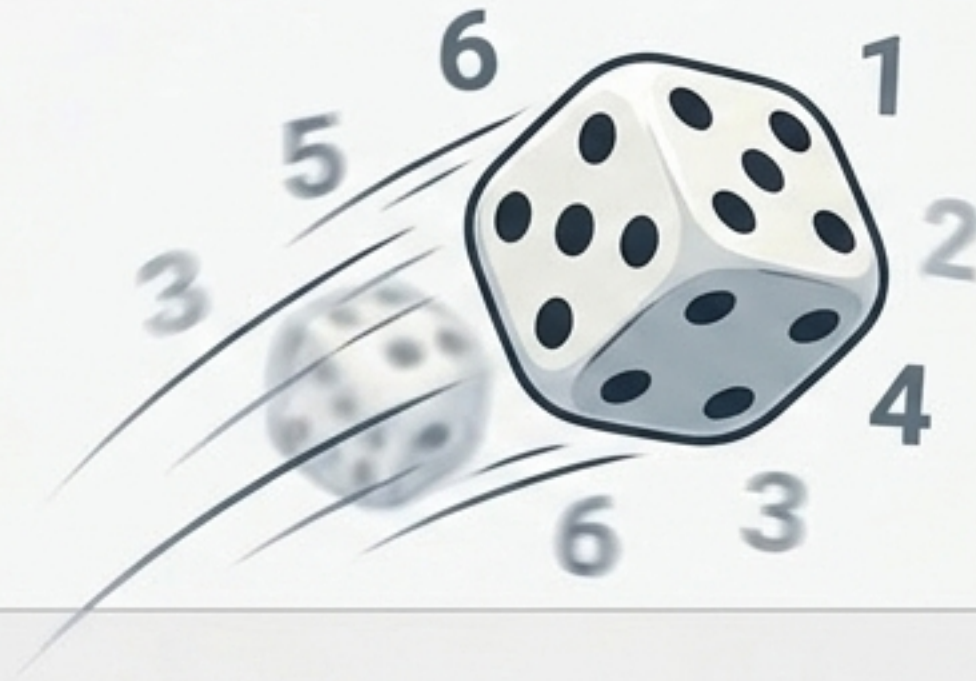
(E) بررسی سریع

سوال: اگر از تمام دانشجویان یک دانشگاه نظرسنجی کنیم، آیا این یک نمونه است؟
پاسخ غلط رایج: بله، نمونه بزرگی است.
پاسخ صحیح: خیر. اگر همه بررسی شوند، ما خود «جامعه» (سرشماری) را بررسی کرده‌ایم، نه یک نمونه.

⚠ اشتباه رایج:

اشتباه گرفتن نمادها: استفاده از n (حجم نمونه) به جای N (حجم کل جامعه) در فرمول‌ها.

متغیر تصادفی در نمونه‌گیری (Random Variables)



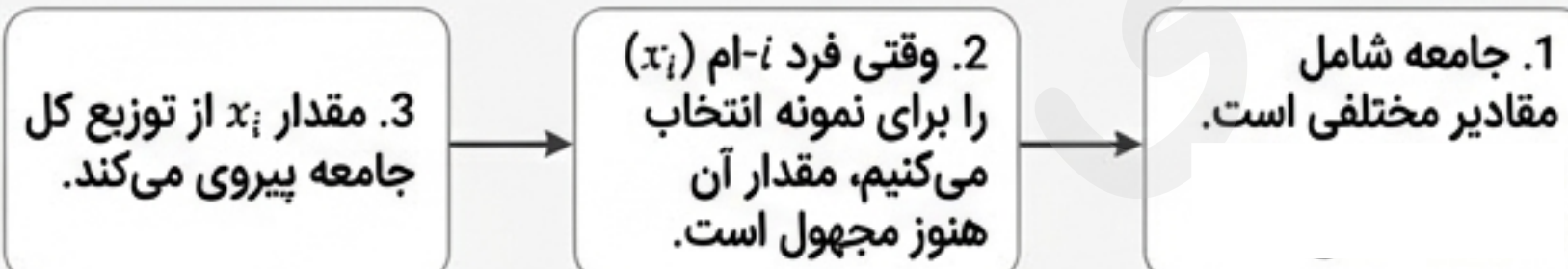
(A) مفهوم

- متغیر تصادفی (x_i): مقدار هر نمونه قبل از استخراج. از آنجا که انتخاب تصادفی است، مقدار آن نامشخص و دارای احتمال است.

(B) شهود

انداختن یک تاس را تصور کنید. قبل از پرتاب، نتیجه یک متغیر تصادفی است که می‌تواند ۱ تا ۶ باشد. پس از پرتاب، تبدیل به یک عدد ثابت می‌شود. نمونه‌گیری نیز دقیقاً همین‌طور است.

(C) نحوه کار



(D) مثال کاربردی

می‌خواهیم قد ۵ نفر را در خیابان بپرسیم. نفر اول (x_i) یک متغیر تصادفی است. وقتی او را متوقف می‌کنیم و قدش ۱۷۵ سانتی‌متر می‌شود، این متغیر به یک مقدار ثابت تبدیل شده است.

(E) بررسی سریع

سوال: آیا داده‌های ثبت‌شده در فایل اکسل شما متغیرهای تصادفی هستند؟

پاسخ غلط رایج: بله، چون به صورت تصادفی جمع‌آوری شده‌اند.
پاسخ صحیح: خیر، آن‌ها «تحقق‌یافته» و مقادیر ثابت‌اند.
متغیر تصادفی مربوط به قبل از مشاهده است.

⚠ اشتباه رایج: فرض اینکه متغیر تصادفی همیشه به معنای یک عدد غیرقابل پیش‌بینی و بی‌نظم است؛ در حالی که از قوانین احتمال توزیع جامعه پیروی می‌کند.

میانگین نمونه (The Sample Mean)



(A) مفهوم

میانگین نمونه ($\bar{x}$): معدل حسابی تمام مقادیر استخراج شده در یک نمونه.

(B) شهود

اگر چند قطب‌نمای ارزان قیمت داشته باشید که هر کدام کمی خطا دارند، گرفتن میانگین اعداد آن‌ها به شما جهت دقیق‌تری نسبت به یک قطب‌نمای منفرد می‌دهد.

(D) مثال کاربردی

برای تخمین حقوق مهندسان، حقوق ۳ نفر (۵۰، ۶۰ و ۷۰) پرسیده می‌شود. میانگین نمونه برابر است با $\bar{x} = 60$.

(C) نحوه کار

1. تمام مقادیر نمونه‌های انتخاب شده (x_i) را جمع کنید.
2. حاصل جمع را بر تعداد نمونه‌ها (n) تقسیم کنید.
3. نتیجه به دست آمده، میانگین نمونه ($\bar{x}$) است.
4. نتیجه به دست آمده، میانگین نمونه ($\bar{x}$) است.

(C.5) کالبدشکافی فرمول

$$\underbrace{\bar{x}}_{\text{نتیجه نهایی میانگین نمونه}} = \underbrace{\frac{1}{n}}_{\text{تقسیم بر اندازه نمونه؛ ثابت معلوم}} \underbrace{\sum}_{\text{عملگر مجموع}} \underbrace{x_i}_{\text{مقدار هر متغیر تصادفی}}$$

بیان ساده: میانگین نمونه برابر است با مجموع همه مقادیر نمونه‌ها تقسیم بر تعداد آن‌ها.
جایگذاری عددی: اگر نمونه‌ها قندهای ۱۷۰، ۱۸۰ و ۱۶۰ باشند:
 $(170+180+160)/3 = 170\$$
فرض پنهان: وزن یا اهمیت تمام اعضای نمونه کاملاً با هم برابر است.

⚠ اشتباه رایج:

اشتباه گرفتن $\bar{x}$ (میانگین نمونه ما) با μ (میانگین واقعی و ثابت کل جامعه).

(E) بررسی سریع

سوال: آیا $\bar{x}$ همیشه دقیقاً برابر با میانگین واقعی جامعه (μ) است؟
پاسخ غلط رایج: بله، چون فرمول میانگین را محاسبه می‌کند.
پاسخ صحیح: خیر. $\bar{x}$ فقط یک «تخمین» متغیر است.

امید ریاضی یک نمونه (Expected Value)

(A) مفهوم

- امید ریاضی ($E[...]$): مقدار متوسطی که در بلندمدت از یک متغیر تصادفی انتظار داریم.
- امید ریاضی هر تک نمونه (x_i) برابر با میانگین کل جامعه (μ) است.

(B) شهود

اگر دست در یک کیسه پر از مهره‌های شماره‌دار ببرید، منطقی‌ترین حدس برای شماره آن مهره، میانگین شماره تمام مهره‌های داخل کیسه است.

(C) نحوه کار

1. جامعه یک توزیع اصلی با میانگین μ دارد.
2. به دلیل انتخاب تصادفی، هیچ نمونه‌ای ارجحیت ندارد.
3. ارزش مورد انتظار هر نمونه جدید به سمت مرکز جامعه میل می‌کند.

⚠ اشتباه رایج:

فرض اینکه یک نمونه به تنهایی می‌تواند نماینده خوبی از کل جامعه باشد.

(D) مثال کاربردی

- میانگین نمره ریاضی کل کشور ۱۵ است. اگر یک دانش‌آموز تصادفی انتخاب کنید، قبل از پرسیدن نمره‌اش، امید ریاضی نمره او ۱۵ است.

(D) مثال کاربردی

میانگین نمره ریاضی کل کشور ۱۵ است. اگر یک دانش‌آموز تصادفی انتخاب کنید، قبل از پرسیدن نمره تصادفی انتخاب کنید، قبل از پرسیدن نمره‌اش، امید ریاضی نمره او ۱۵ است.



(C.5) کالبدشکافی فرمول

$$E[x_i] = \mu$$

میانگین نامعلوم جامعه واقعی
متغیر تصادفی i -ام

بیان ساده: حدس میانگین ما برای هر عضو جدید نمونه، همان میانگین کل جامعه است.

جایگذاری عددی: اگر میانگین قد جامعه ۱۷۵ باشد ($\mu = 175$)، حدس ما برای نفر بعدی ۱۷۵ است: $E[x_2] = 175$
فرض پنهان: نمونه‌گیری کاملاً تصادفی و بدون سوگیری است.

(E) بررسی سریع

سوال: آیا $E[x_i] = \mu$ به این معناست که مقدار نمونه حتماً همان μ خواهد بود؟
پاسخ غلط رایج: بله، ریاضیات می‌گویند باید برابر باشد.

پاسخ صحیح: خیر. این فقط یک میانگین احتمالی نظری است، نه یک مقدار قطعی.

امید ریاضی میانگین نمونه (Expected Value of Sample Mean)

(A) مفهوم

- تخمین‌گر بدون سوگیری (Estimator Unbiased): فرمولی که در میان‌مدت خطای سیستماتیک ندارد.
- میانگین نمونه $(\bar{x})$ ، یک تخمین‌گر بدون سوگیری برای میانگین جامعه (μ) است.

(B) شهود

تیراندازی را تصور کنید که گاهی به چپ و گاهی به راست مرکز هدف می‌زند. «مرکز تجمع» تیرهای او دقیقاً روی هدف اصلی است.

(E) بررسی سریع

سوال: کلمه «بدون سوگیری» (Unbiased) در آمار چه معنایی دارد؟

پاسخ غلط رایج: یعنی تخمین ما هیچ خطایی ندارد.

پاسخ صحیح: یعنی خطای سیستماتیک ندارد؛ خطاها یکدیگر را خنثی می‌کنند و معدل خطا صفر است.

۰۹۳۸۳۹۴۴۴۷۷



(D) مثال کاربردی

شما ۱۰۰ بار نمونه‌های ۱۰ تایی از لامپ‌ها می‌گیرید. میانگین این ۱۰۰ میانگین متفاوت، به شدت به عمر واقعی لامپ‌های کارخانه نزدیک است.

(C) نحوه کار

1. قانون مجموع‌ها: $E[A + B] = E[A] + E[B]$.
2. میانگین نمونه مجموعی از x_i ها تقسیم بر n است.
3. چون هر $E[x_i] = \mu$ ، پس میانگین n تا μ نیز برابر با خود μ خواهد شد.

(C.5) کالبدشکافی فرمول

$$E[\bar{x}] = \mu$$

میانگین جامعه امید ریاضی میانگین نمونه

بیان ساده: در تکرار زیاد، میانگین، میانگین‌های شما، همان میانگین واقعی جامعه خواهد شد.

جایگذاری عددی: $E\left[\frac{\mu + \mu + \dots}{n}\right] = \frac{n \times \mu}{n} = \mu$

فرض پنهان: جامعه تغییر نمی‌کند و نمونه‌گیری کاملاً تصادفی است.

⚠ اشتباه رایج:

باور به اینکه «بدون سوگیری» بودن میانگین نمونه، ضامن دقیق بودن آن در یک نمونه‌گیری خاص است.

چوادرودخی

واریانس میانگین نمونه: مفهوم (Variance of Sample Mean)

(A) مفهوم

- واریانس میانگین نمونه: معیاری است که نشان می‌دهد مقادیر $\bar{x}$ در نمونه‌گیری‌های مختلف چقدر از یکدیگر و از μ پراکنده‌اند.

(B) شهود

تصور کنید چند تیم نقشه‌برداری به یک تپه می‌روند. اگر دستگاه‌هایشان دقیق باشد، گزارش گزارش‌هایشان بسیار شبیه هم است (واریانس کم). دستگاه‌ها خراب باشد، هر تیم عدد متفاوتی می‌دهد (واریانس زیاد).

(C) نحوه کار

1. هر بار که یک نمونه می‌گیرید، یک $\bar{x}$ متفاوت به دست می‌آید.
2. این $\bar{x}$ ها برای خودشان یک توزیع تشکیل می‌دهند.
3. میزان گستردگی این توزیع، میزان اطمینان ما را تعیین می‌کند.

Low Variance (دقت بالا، واریانس کم)



High Variance (دقت پایین، واریانس زیاد)



(D) مثال کاربردی

اگر در نظرسنجی انتخابات، پیش‌بینی شود کاندیدای (الف) ۵۰٪ رای می‌آورد، واریانس مشخص می‌کند که این تخمین چقدر «پهن» است (بین ۴۰٪ تا ۶۰٪ یا دقیقاً ۴۹٪ تا ۵۱٪).

(E) بررسی سریع

سوال: آیا واریانس میانگین نمونه با واریانس خود جامعه برابر است؟

پاسخ غلط رایج: بله، چون از همان جامعه گرفته شده است.

پاسخ صحیح: خیر! میانگین‌ها همیشه پراکندگی بسیار کمتری نسبت به تک‌تک افراد جامعه دارند.

⚠ اشتباه رایج:

خلط کردن «پراکندگی افراد در یک جامعه» با «پراکندگی میانگین نمونه‌ها در آزمون‌های مختلف».

فرض استقلال (Assumption of Independence)

(A) مفهوم

- استقلال در آمار: وقوع یک اتفاق (یا استخراج یک نمونه) هیچ تاثیری بر نتیجه استخراج نمونه بعدی نداشته باشد.

(B) شهود

برداشتن قطره آب از اقیانوس، تاثیری بر ترکیب قطره بعدی نمی‌گذارد (مستقل).
اما برداشتن آب‌نبات قرمز از شیشه ده تایی، شانس آب‌نبات قرمز بعدی را کم می‌کند (وابسته).



(C) نحوه کار

1. وقتی جامعه بسیار بزرگ است ($N \rightarrow \infty$)، خروج یک عضو بی‌تاثیر است.
2. بنابراین نمونه‌ها از یکدیگر مستقل فرض می‌شوند.
3. در حالت استقلال، واریانس مجموع چند متغیر، برابر با مجموع واریانس‌های آن‌هاست.

(D) مثال کاربردی

نظرسنجی از ۱۰۰۰ نفر در کشوری با ۸۰ میلیون جمعیت. انتخاب نفر اول تغییری در شانس انتخاب ۸۰ میلیون نفر باقی‌مانده ایجاد نمی‌کند. پس نظرات مستقل فرض می‌شوند.

(E) بررسی سریع

سوال: در چه شرایطی فرض استقلال نقض می‌شود؟
پاسخ غلط رایج: وقتی افراد نمونه با هم در یک اتاق باشند.
پاسخ صحیح: وقتی حجم جامعه نسبتاً کوچک است و اعضا پس از انتخاب به جامعه برگردانده نمی‌شوند.

⚠ اشتباه رایج:

استفاده کورکورانه از فرمول‌های ساده آماری (که فرض استقلال دارند) برای جوامع و کلامهای کوچک (مثلاً ۲۰ نفره).

واریانس میانگین نمونه (حالت تقریب)

(A) مفهوم

برای یک جامعه بسیار بزرگ (استقلال برقرار است)، واریانس میانگین نمونه برابر است با واریانس جامعه تقسیم بر اندازه نمونه.

(B) شهود

هرچه تیم نظرسنجی شما بزرگتر باشد (n بزرگتر)، خطای میانگین نظرات آن‌ها کمتر می‌شود. قدرت جمع، واریانس فردی را خنثی می‌کند.

(C) نحوه کار

1. واریانس ثابت ($1/n$) در متغیر (X) برابر است با $(1/n^2) \times \text{Var}(X)$.
2. مجموع واریانس‌های n نمونه برابر با $n \times \sigma^2$ است.
3. ضرب این دو: $(1/n^2) \times (n \times \sigma^2) = \sigma^2/n$.

(D) مثال کاربردی

واریانس وزن نوزادان ۵۰ است. اگر نمونه‌های ۱۰ تایی بگیریم، واریانس میانگین وزن این گروه‌ها ۵ خواهد بود ($\$50/10$).

(E) بررسی سریع

سوال: اگر اندازه نمونه (n) را دو برابر کنیم، واریانس چه تغییری می‌کند؟
پاسخ غلط رایج: انحراف معیار نصف می‌شود.
پاسخ صحیح: واریانس نصف می‌شود.

⚠ اشتباه رایج: فراموش کردن به توان دو رساندن مقدار مخرج ($1/n$) هنگام خروج آن از عملگر واریانس.



(C.5) کالبدشکافی فرمول

واریانس اصلی جامعه

$$\text{Var}(\bar{x}) \approx \sigma^2 / n$$

واریانس میانگین نمونه

تقریباً برابر است با

تقسیم بر اندازه نمونه

بیان ساده: پراکندگی میانگین نمونه‌ها، دقیقاً به نسبت اندازه نمونه کاهش می‌یابد. جایگذاری عددی: اگر $\sigma^2 = 100$ و $n = 25$ باشد، واریانس میانگین برابر $100/25 = 4$ خواهد شد. فرض پنهان: جامعه بی‌نهایت بزرگ است و نمونه‌ها مستقل از یکدیگرند.

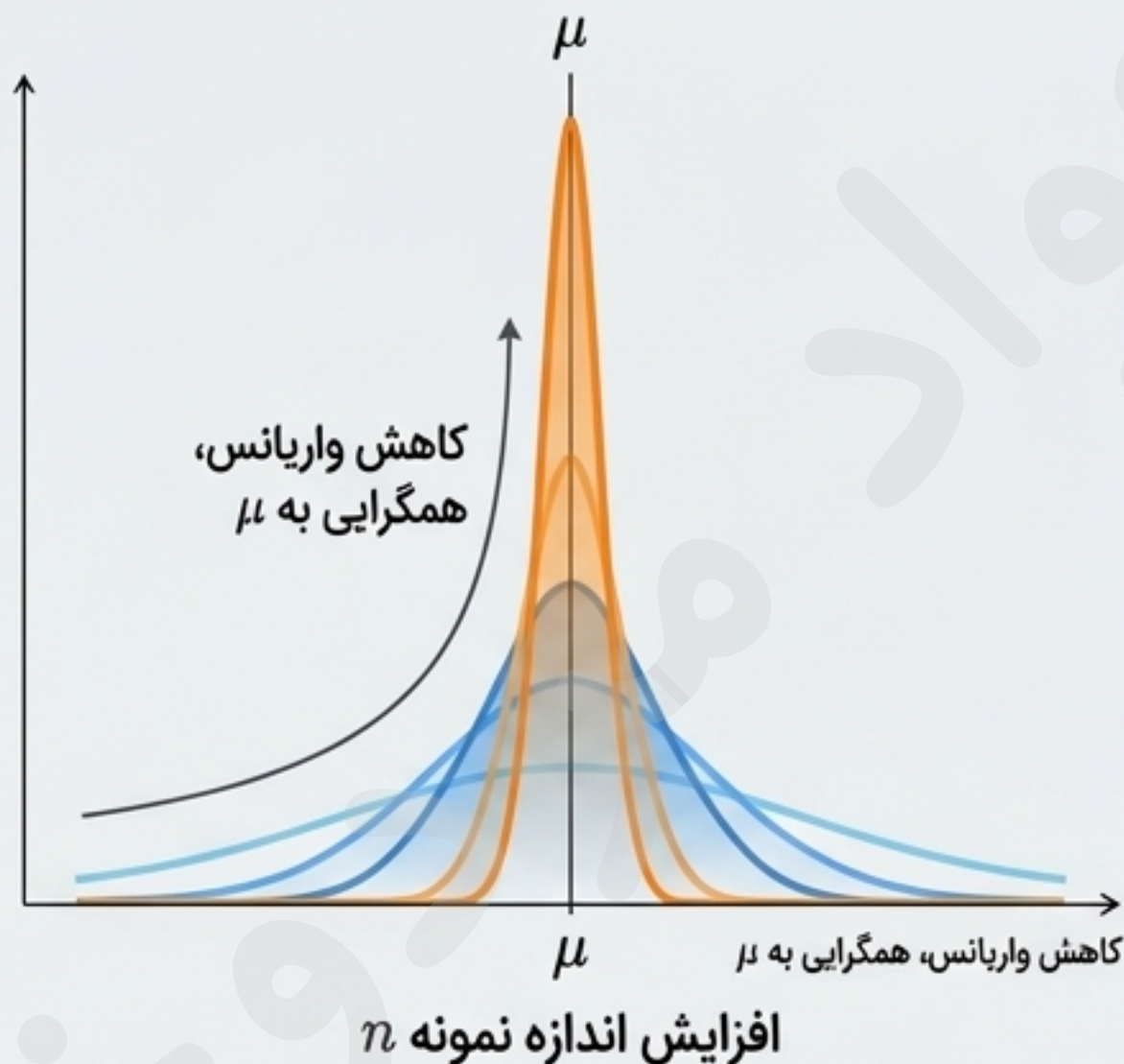
تاثیر اندازه نمونه (Law of Large Numbers)

(A) مفهوم

- قانون اعداد بزرگ: با افزایش اندازه نمونه (n)، واریانس کاهش یافته و میانگین نمونه ($\bar{x}$) به میانگین جامعه (μ) همگرا می‌شود.

(B) شهود

اگر سکه را ۳ بار بیندازید، ممکن است هر ۳ بار شیر بیاید (خطای زیاد).
اگر ۳۰۰۰ بار بیندازید، محال است به چیزی جز حدود ۵۰٪ برسید. تعداد بالا، نویزها را می‌شوید.



(C) نحوه کار

1. بر اساس فرمول $\frac{\sigma^2}{n}$ ، با میل کردن n به سمت بی‌نهایت، کسر به سمت صفر میل می‌کند.
2. واریانس صفر یعنی هیچ پراکندگی وجود ندارد.
3. توزیع آماری به یک خط عمودی دقیق روی μ تبدیل می‌شود.

(D) مثال کاربردی

در یادگیری ماشین (Machine Learning)، آموزش مدل با ۱۰۰ عکس گربه، خطای تشخیص بالایی دارد. با افزایش به ۱,۰۰۰,۰۰۰ عکس، مدل به توانایی تشخیص دقیق دست می‌یابد.

(E) بررسی سریع

سوال: آیا برای دوبرابر کردن دقت تخمین، کافی است اندازه نمونه را دو برابر کنیم؟
پاسخ غلط رایج: بله، رابطه مستقیم دارند.
پاسخ صحیح: خیر، چون انحراف معیار (دقت) با رادیکال n در ارتباط است. برای دوبرابر شدن دقت، باید نمونه ۴ برابر شود!

⚠ اشتباه رایج:

تصور اینکه جمع‌آوری دیتای بیشتر همیشه مشکلات مدل را خطی حل می‌کند (نادیده گرفتن قانون بازده نزولی).

نمونه‌گیری بدون جایگذاری (Without Replacement)

(A) مفهوم

- نمونه‌گیری بدون جایگذاری: استخراج نمونه‌ها به گونه‌ای که هر فرد انتخاب‌شده، دوباره به جامعه برنمی‌گردد. (حجم جامعه پس از هر استخراج کوچکتر می‌شود).

(B) شهود

وقتی کارت آس پیک به کسی داده می‌شود، احتمال دریافت آن برای نفر بعدی صفر می‌شود! انتخاب اول، سرنوشت انتخاب دوم را تغییر می‌دهد.

(C) نحوه کار

1. حجم اولیه جامعه N است.
2. نمونه اول استخراج می‌شود. حالا جامعه $N-1$ عضو دارد.
3. نمونه دوم استخراج می‌شود. جامعه $N-2$ عضو دارد.
4. این کاهش باعث تغییر مداوم در احتمالات برای نمونه‌های بعدی می‌شود.

(D) مثال کاربردی

مدیر یک شرکت ۵۰ نفره می‌خواهد ۳ نفر را اخراج کند. وقتی نفر اول انتخاب می‌شود، استخر باقی‌مانده ۴۹ نفر است. استخراج‌ها کاملاً به یکدیگر متصل و وابسته‌اند.

(E) بررسی سریع

سوال: در نظرسنجی‌های تلفنی عمومی، نمونه‌گیری با جایگذاری است یا بدون جایگذاری؟
پاسخ غلط رایج: با جایگذاری، چون جامعه بزرگ است.
پاسخ صحیح: بدون جایگذاری. ما هیچ‌گاه به یک فرد دو بار زنگ نمی‌زنیم. (اما در جامعه بزرگ، شبیه با جایگذاری عمل می‌کند).

⚠ اشتباه رایج:

فرض اینکه در تمام مسائل آمار، نمونه‌ها مستقل هستند. در نمونه‌گیری بدون جایگذاری، متغیرها کاملاً «وابسته» (Dependent) می‌شوند.



وابستگی و کوواریانس (Dependence & Covariance)

(A) مفهوم

- کوواریانس (Covariance): معیاری برای اندازه‌گیری نحوه تغییر دو متغیر تصادفی در کنار یکدیگر.
- در نمونه‌گیری بدون جایگذاری، بین نمونه‌ها یک کوواریانس کوچک ایجاد می‌شود.

(C) نحوه کار

1. اگر x_i (نمونه اول) مقداری بزرگتر از میانگین داشته باشد و خارج شود...
2. توزیع مقادیر باقی‌مانده به سمت اعداد کوچکتر متمایل می‌شود.
3. در نتیجه، x_j (نمونه بعدی) تمایل به داشتن مقدار کوچکتری پیدا می‌کند (حرکت در جهت عکس).

⚠ اشتباه رایج:

نادیده گرفتن کوواریانس به دلیل سخت بودن محاسبه آن، که در جوامع کوچک منجر به تخمین اشتباه اشتباه خطای اندازه‌گیری می‌شود.

(B) شهود

الاکلنگ را در نظر بگیرید. اگر یک طرف بالا برود، طرف دیگر مجبور است پایین بیاید (کوواریانس منفی). خروج فرد بسیار بلند از جامعه، میانگین قد افراد باقی‌مانده را کمی افت می‌دهد.

(D) مثال کاربردی

قرعه‌کشی با ۱۰ توپ (۵ قرمز، ۵ آبی).
اگر توپ اول قرمز باشد، توپ‌های باقی‌مانده (۴ قرمز، ۵ آبی) احتمال بیشتری برای آبی بودن توپ دوم ایجاد می‌کنند.
این کوواریانس منفی است.

(E) بررسی سریع

سوال: اگر نمونه‌ها مستقل باشند، کوواریانس آن‌ها چقدر است؟
پاسخ غلط رایج: یک عدد بسیار کوچک.
پاسخ صحیح: دقیقاً صفر. وابستگی صفر مساوی است با کوواریانس صفر.



فرمول کوواریانس نمونه‌ها (The Covariance Lemma)

(A) مفهوم

- قضیه کوواریانس: بین هر دو نمونه متمایز در نمونه‌گیری بدون جایگذاری، یک کوواریانس منفی ثابت وجود دارد.

(B) شهود

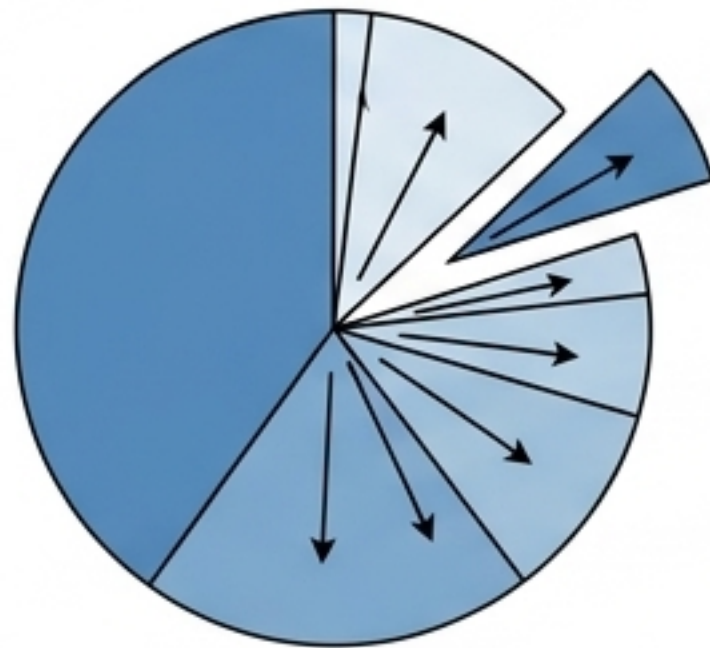
کشیدن یک فرد از جامعه، «سهم» بقیه را کمی تغییر می‌دهد. این تغییر به اندازه واریانس جامعه است که مساوی بین افراد باقی‌مانده ($N - 1$) تقسیم و از باقی‌مانده و تقسیم و از شانس آن‌ها کم شده است.

(C) نحوه کار

1. متناسب با واریانس کل جامعه (σ^2) است.
2. چون با برداشتن یک عضو ایجاد می‌شود، علامت آن منفی است.
3. اثر این برداشت، بین $N - 1$ عضو باقی‌مانده تقسیم می‌شود.

(C.5) کالبدشکافی فرمول

بیان ساده: هر دو نمونه با هم یک رابطه عکس دارند که مقدار آن به بزرگی جامعه بستگی دارد. جایگذاری عددی: اگر $\sigma^2 = 100$ و $N = 101$ باشد، کوواریانس برابر $-100/(100 = -1)$ خواهد بود. فرض پنهان: اندیس‌های دو نمونه با هم برابر نیستند ($i \neq j$).



$$\underbrace{Cov(x_i, x_j)}_{\text{کوواریانس بین دو نمونه متمایز}} = \underbrace{-}_{\text{علامت منفی ثابت}} \underbrace{\sigma^2}_{\text{واریانس اصلی جامعه}} / \underbrace{(N - 1)}_{\text{حجم جامعه منهای یک}}$$

⚠ اشتباه رایج:

مثبت در نظر گرفتن مقدار این کوواریانس. به دلیل ماهیت بدون جایگذاری، همواره منفی است.

(E) بررسی سریع

سوال: با بزرگ شدن بی‌نهایت جامعه چه می‌شود؟ پاسخ صحیح: به سمت صفر میل می‌کند! (دلیل تقریب مستقل فرض کردن در جوامع بزرگ).

(D) مثال کاربردی

کلاس ۲۰ نفره ($N = 20$) با واریانس نمرات ۳۸. کوواریانس نفر اول و دوم: $-38 / 19 = -2$.

گسترش فرمول واریانس مجموع (Variance of a Sum)

(A) مفهوم

وقتی متغیرها مستقل نیستند، واریانس مجموع آن‌ها فقط جمع ساده‌ی واریانس‌هایشان نیست، بلکه باید تمامی تعاملات دوگانه (کوواریانس‌ها) را هم وارد فرمول کنیم.

(B) شهود

بسته‌بندی وسایل شکننده. اگر لیوان‌ها (واریانس فردی) را کنار هم بگذارید، احتمال شکستن کل جعبه به اینکه چقدر لیوان‌ها به هم برخورد می‌کنند (کوواریانس) نیز بستگی دارد.

(D) مثال کاربردی

تخمین ریسک سبد سهام. ریسک کل سبد فقط جمع ریسک تک‌تک سهام‌ها نیست، بلکه به اینکه چقدر این سهام‌ها با هم سقوط یا صعود می‌کنند به شدت وابسته است.

(E) بررسی سریع

سوال: در فرمول بسط یافته، چند جفت کوواریانس مختلف برای یک نمونه ۳ تایی وجود دارد؟
پاسخ غلط رایج: ۳ جفت.
پاسخ صحیح: ۶ جفت.
طبق رابطه $n(n-1)$ ، داریم $3(2) = 6$ ترکیب دوگانه.

(C) نحوه کار

۱. واریانس میانگین معادل واریانس مجموع مقادیر ضرب در ثابت $1/n^2$ است.
۲. بسط دادن واریانس مجموع، به ماتریسی از واریانس‌ها و کوواریانس‌ها تبدیل می‌شود.
۳. ما n تا واریانس و $n(n-1)$ تا کوواریانس خواهیم داشت.



⚠ اشتباه رایج:

فراوانی اشتباه: دانشجویان اغلب تصور می‌کنند فقط n تا کوواریانس وجود دارد، نه $n(n-1)$ تا.

ضریب تصحیح جامعه متناهی (Finite Population Correction (FPC)

(A) مفهوم

- ضریب تصحیح (Finite Population Correction):
یک کسر تعدیل کننده است که خطای فرض استقلال در جوامع کوچک را جبران و اصلاح می کند.

(B) شهود

خالی کردن بادکنک.
سوزن کوچک = تخلیه آرام.
سوراخ بزرگ (نمونه بزرگ نسبت به جامعه) = افت فشار ناگهانی.
FPC این «تغییر فشار» را محاسبه می کند.

(C) نحوه کار

1. با جابگذاری ترم های کوواریانس و فاکتورگیری.
2. یک کسر جدید در فرمول ظاهر می شود: $(N-n)/(N-1)$.
3. هرچه نمونه بخش بزرگ تری از جامعه را بپوشاند، این ضریب عدد کوچکتری شده و واریانس را بپوشاند، این ضریب عدد کوچکتری شده و واریانس را به شدت کاهش می دهد.

(D) مثال کاربردی

- اگر در شرکت ۱۰ نفره از ۹ نفر نظرسنجی کنید، شما تقریباً نظر کل جامعه را پرسیده اید. ضریب $(1/9 \approx 0.11)$ خطای فرمول های استاندارد را تصحیح می کند.



سوزن کوچک = تخلیه آرام

سوراخ بزرگ (نمونه بزرگ نسبت به جامعه) = افت فشار ناگهانی

(C.5) کالبدشکافی فرمول

ضریب تصحیح

$$FPC = \frac{\text{حجم جامعه منهای حجم نمونه } (N - n)}{\text{حجم جامعه منهای یک } (N - 1)}$$

بیان ساده: اگر اندازه نمونه شما نزدیک به کل جامعه شود، خطای تخمین شما به شدت کاهش می یابد. جایگذاری عددی: جامعه ۱۰۰ نفره و نمونه ۴۰ نفره. ضریب برابر $(100-40)/99$ (واریانس ۴۰٪ کم می شود!). $0.60 \approx (10 - 40) = 10 - 10$ (واریانس ۱۰ - ۱۰) فرض پنهان: نمونه گیری بدون جایگذاری است.

سوال: اگر از کل جامعه نمونه بگیریم ($n = N$)، مقدار FPC چقدر می شود؟
پاسخ صحیح: صفر! یعنی هیچ خطایی در تخمین نخواهیم داشت.

⚠ اشتباه رایج:

ضرب نکردن این ضریب در مواقعی که حجم نمونه بیش از ۵٪ حجم کل جامعه است.

فرمول دقیق واریانس میانگین (Exact Variance)

(A) مفهوم

- واریانس دقیق میانگین نمونه: ترکیب فرمول تقریب (برای جوامع بی‌نهایت) ضربدر ضریب FPC (برای اصلاح کاهش تنوع در جامعه محدود).

(B) شهود

مانند خرید با کد تخفیف. فرمول اصلی، قیمت پایه است. ضریب FPC، همان کد تخفیفی است که وقتی حجم خریدتان بالا می‌رود، مبلغ پرداختی (خطا) را پایین می‌آورد.

(D) مثال کاربردی

از ۵۰۰ سند (N)، یک حسابرس ۲۰۰ سند (n) را چک می‌کند (بدون تکرار). چون کسر ۴۰٪ است، واریانس باید حتماً در ضریب ضرب شود.

(C) نحوه کار

1. واریانس با فرض استقلال به دست می‌آید: σ^2/n
2. ضریب تصحیح محاسبه می‌شود: $(N - n)/(N - 1)$
3. این دو در هم ضرب می‌شوند.

C.5) کالبدشکافی فرمول

بیان ساده: واریانس محدود، برابر است با فرمول ساده که توسط ضریب ابعاد جامعه، متراکم شده است.
جایگذاری عددی: پایه=۵۰، $N = 100$ ، $N = 100$.
مقدار دقیق: $(50) \times (90/99) = 45.45$.
فرض پنهان: مقدار N مشخص و دقیق است.

(E) بررسی سریع

سوال: چرا همیشه از فرمول دقیق استفاده نمی‌کنیم؟
پاسخ صحیح: زیرا در بسیاری از مواقع، اندازه واقعی جامعه (N) بی‌نهایت یا کاملاً نامشخص است!

قیمت پایه (Base Price)

$$\sigma^2/n$$

کد تخفیف: FPC

\$0.45 × — ضریب تصحیح
جامعه مناهی
(Multiplier applied)

جمع نهایی (Final Total)

$$Var(\bar{x}) = \frac{\sigma^2}{n} \times \frac{N - n}{N - 1}$$

— واریانس دقیق
(Exact Variance)

$$Var(\bar{x}) = \left[\frac{\sigma^2}{n} \right] \times \left[\frac{(N - n)}{(N - 1)} \right]$$

واریانس دقیق / میانگین نمونه
 بخش تقریب کلاسیک / ضرب در
 ضریب تصحیح جامعه متناهی - FPC

⚠ اشتباه رایج: استفاده از این فرمول در توزیع‌های «با جایگذاری». اگر نمونه به جامعه برگردد، FPC بی‌معناست.

چه زمانی از کدام فرمول استفاده کنیم؟ (Decision Rule)

(A) مفهوم

یک قانون سرانگشتی (Rule of Thumb) برای انتخاب بین استفاده از «تقریب» یا «فرمول دقیق با تصحیح FPC».

(B) شهود

اگر نمونه شما در برابر جامعه ناچیز است (پشه روی فیل)، از اثر آن چشم‌پوشی کنید. اما اگر نمونه وزن قابل توجهی دارد (پرندۀ روی گربه)، محاسبات را اصلاح کنید.

(C) نحوه کار (فلوچارت تصمیم)

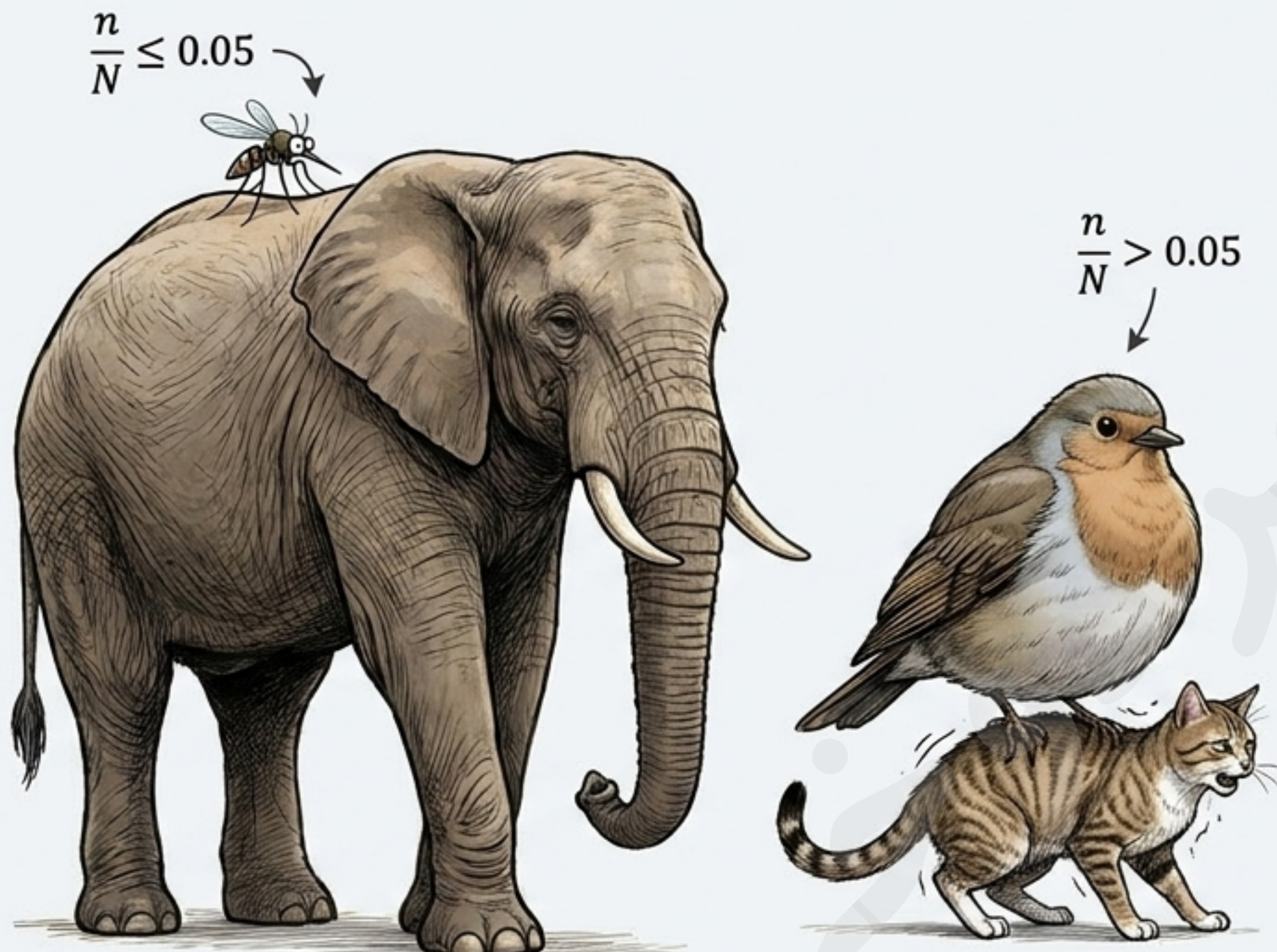
1. کسر n/N را محاسبه کنید.
2. اگر $n/N \leq 0.05$ (کمتر از ۵٪) → از تقریب σ^2/n استفاده کنید.
3. اگر $n/N > 0.05$ (بیشتر از ۵٪) → باید حتماً از فرمول دقیق با FPC استفاده کنید.

(D) مثال کاربردی

- نمونه ۱,۰۰۰ نفری از ۹ میلیون شهروند (۰,۰۱٪) → بدون FPC.
- نمونه ۵۰ نفری از دانشکده ۲۰۰ نفری (۲۵٪) → استفاده از تصحیح قطعی است.

(E) بررسی سریع

سوال: اگر جامعه نامشخص اما بسیار بزرگ باشد، چه کنیم؟
پاسخ غلط رایج: نمی‌توانیم واریانس را حساب کنیم.
پاسخ صحیح: وقتی N به قدری بزرگ است که n/N به صفر میل می‌کند، ترم FPC عملاً ۱ می‌شود. با خیال راحت تقریب را به کار می‌بریم.

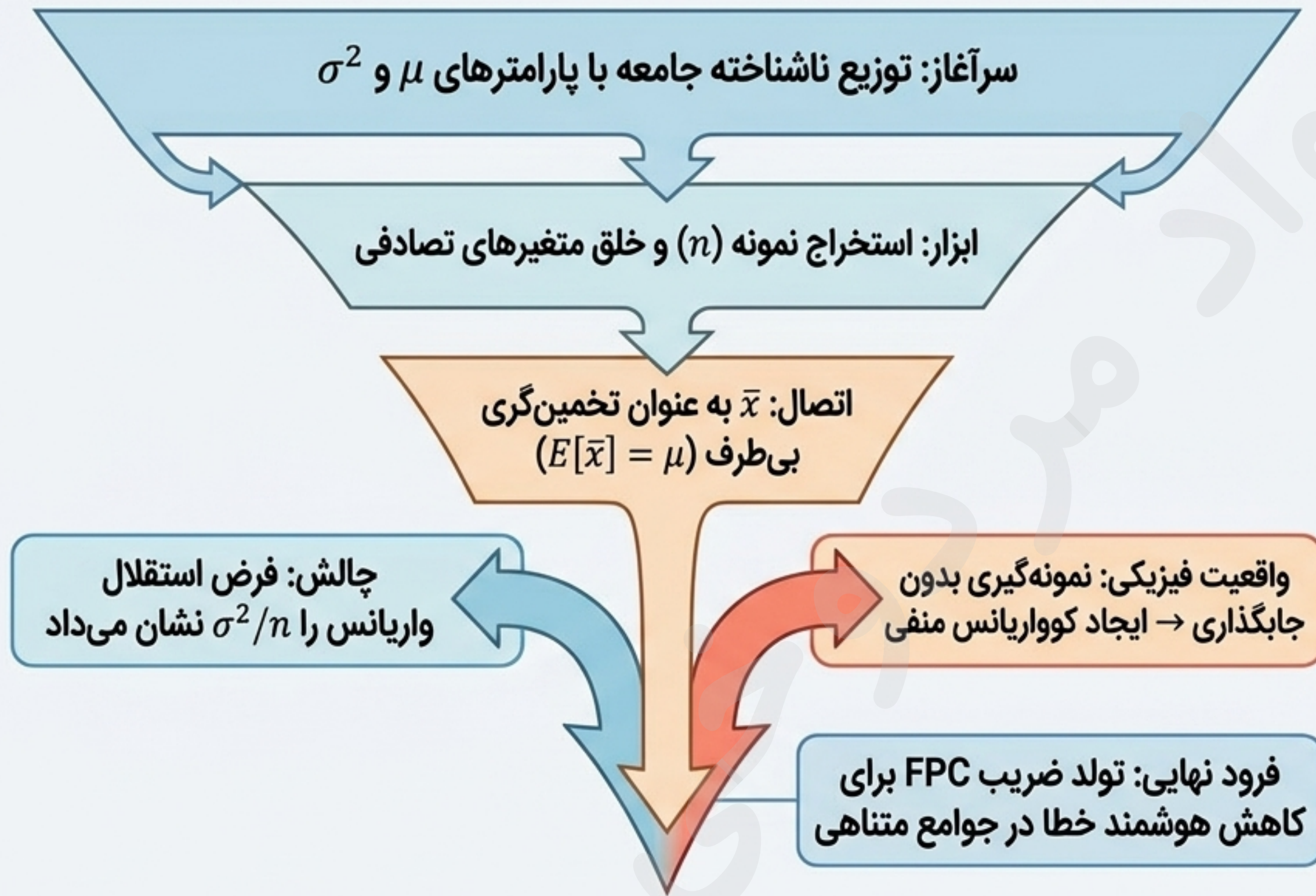


اثر ناچیز

وزن قابل توجه

⚠️ **اشتباه رایج:** اصرار روی استفاده از فرمول‌های ساده در تحقیقات میدانی با جوامع محدود شرکتی که خطای فاحشی در گزارش تخمین ایجاد می‌کند.

ترکیب و پیوند مفاهیم (Synthesis: The Statistical Lens)



۳ بینش کلان (Big Picture Insights)

1. استقلال یک توهم مفید است: در دنیای دنیای واقعی داده‌های بدون جایگذاری مستقل نیستند، اما در اعداد بزرگ اثر آن قابل چشم‌پوشی است.
2. ریاضیات خوداصلاح‌گر است: وقتی نمونه نمونه به اندازه خود جامعه برسد، فرمول FPC به صورت هوشمند خطا را صفر می‌کند!
3. اندازه نسبی اهمیت دارد: قدرت یک نمونه فقط به تعداد آن (n) نیست، بلکه به نسبت آن در جامعه (n/N) وابسته است.

⚠ اشتباه کلان (The Big Mistake): گم شدن در فرمول‌های جبری و فراموش کردن هدف اصلی: «چقدر می‌توانم به میانگینی که روی کاغذ حساب کرده‌ام، اعتماد کنم؟»

نکات کلیدی و تله‌های رایج (Takeaways & Pitfalls)

نکات کلیدی (Key Takeaways):

- ✓ یک تک‌نمونه تصادفی ذاتاً متغیر است و امید ریاضی‌اش μ است.
- ✓ $\bar{x}$ یک تخمین‌گر کاملاً بی‌طرف (Unbiased) از جامعه است.
- ✓ واریانس کمتر = دقت بیشتر در تخمین.
- ✓ نمونه‌گیری بدون جایگذاری، استقلال را نقض می‌کند.
- ✓ خروج یک عضو، شانس اعضای باقیمانده را برهم می‌زند (کوواریانس منفی).
- ✓ ضریب FPC برای جبران این اثر در جوامع کوچک استفاده می‌شود.



۳ تله رایج (Top 3 Pitfalls):

- تله نمادها: جابجا گرفتن $\bar{x}$ (دیتای ما) با μ (حقیقت پنهان).
- تله استقلال: تصور اینکه هر رندوم سمپلینگ به معنای استقلال کامل است.
- تله مخرج: فراموش کردن تفاوت n با N .

🔦 قوانین طلایی (Golden Rules):

- جامعه نامعلوم است؟ $FPC \rightarrow$ را رها کن.
- اگر $n/N > 0.05$ است؟ $\rightarrow$ استفاده از FPC اجباری است.
- امید ریاضی x_i همیشه μ است.



⚠️ اشتباه رایج در مرور: حفظ کردن فرمول‌ها بدون درک بصری رخدادهای فیزیکی در دنیای واقعی.

ارزیابی نهایی و مسیر مطالعه (Final Assessment)

1. امید ریاضی میانگین نمونه به اندازه n بستگی دارد؟

→ (پاسخ: خیر، همیشه μ است)

2. چرا بدون جایگذاری، کوواریانس منفی است؟

→ (پاسخ: چون خروج نمونه، فضای باقیمانده را کاهش می‌دهد)

3. ضریب FPC چه زمانی صفر می‌شود؟

→ (پاسخ: وقتی $n = N$)

4. تخمین‌گر بدون سوگیری یعنی چه؟

→ (پاسخ: در درازمدت خطای سیستماتیک ندارد)

5. قانون ۵٪ چیست؟

→ (پاسخ: اگر n کمتر از ۵٪ جامعه باشد، از FPC صرف‌نظر می‌کنیم)

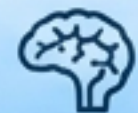
6. محاسبه سریع: $n = 25, \sigma^2 = 100$ در

در جامعه بی‌نهایت، واریانس $\bar{x}$ ؟

→ (پاسخ: ۴)

مسیر مطالعه بعدی (Next Steps):

قضیه حد مرکزی (CLT):
اینکه چگونه علیرغم توزیع عجیب جامعه، میانگین نمونه توزیع زنگوله‌ای پیدا می‌کند.



فواصل اطمینان (Confidence Intervals):
ساخت بازه ۹۵ درصدی با استفاده از واریانس میانگین.



تخمین واریانس جامعه (s^2):
وقتی σ^2 واقعی را نداریم، چگونه آن را با داده‌ها تخمین بزنیم.



⚠️ **اشتباه رایج در یادگیری:** توقف یادگیری در اینجا و تمرین نکردن مباحث با داده‌های واقعی در Python یا Excel.

نسخه کامل این فایل آموزشی را دریافت کنید

شما در حال مشاهده پیش‌نمایش 24 صفحه از مجموع 24 صفحه هستید.
برای دسترسی به فایل کامل به لینک اختصاصی زیر مراجعه کنید.



خرید مستقیم از طریق سایت



خرید از طریق ربات تلگرام



پشتیبانی

تولید محتوا: جواد مردوخی | ریاضیات مفهومی