الف) چگونه قضیه حد مرکزی (CLT) به ما اجازه می‌دهد از نمونه‌های کوچک به جامعه نتیجه‌گیری کنیم؟

قضیه حد مرکزی بیان می‌کند که اگر تعداد نمونه‌ها به اندازه کافی زیاد باشد، توزیع میانگین نمونه‌ها به توزیع نرمال نزدیک می‌شود، حتی اگر توزیع اصلی جامعه نرمال نباشد. این ویژگی باعث می‌شود:

* بتوانیم با استفاده از میانگین یک نمونه، تخمینی از میانگین کل جامعه ارائه کنیم.
* فاصله اطمینان (confidence interval) بسازیم و سطح عدم قطعیت را مشخص کنیم.
* آزمون‌های آماری (مثل Z-test یا t-test) را برای بررسی فرضیات اجرا کنیم.

مثال واقعی:  
در نظرسنجی‌های انتخاباتی، معمولاً با نمونه‌گیری از چند صد یا چند هزار نفر، رفتار رای‌دهندگان یک کشور با میلیون‌ها نفر را تخمین می‌زنند. CLT این امکان را فراهم می‌کند چون می‌دانیم توزیع میانگین‌های نمونه‌ای تقریباً نرمال خواهد بود و می‌توان احتمال خطا را محاسبه کرد.

ب) تأثیر خطاهای نمونه‌گیری بر کاربرد CLT:

اگر نمونه‌گیری به‌درستی انجام نشود، استنتاج‌های مبتنی بر CLT ممکن است گمراه‌کننده باشند. دلایل:

1. سوگیری نمونه (Sampling Bias):  
   اگر نمونه به‌طور تصادفی انتخاب نشود یا بخش خاصی از جامعه در نمونه بیش‌نمایی یا کم‌نمایی شود، میانگین نمونه نماینده خوبی از میانگین جامعه نخواهد بود و کاربرد CLT نادرست خواهد شد.

مثال: اگر فقط از دانش‌آموزان مدارس شهری نمونه‌گیری کنیم و درباره کل کشور نظر بدهیم.

1. اندازه ناکافی نمونه (Small Sample Size):  
   اگر حجم نمونه بسیار کوچک باشد (مثلاً کمتر از 30)، توزیع میانگین نمونه‌ها ممکن است هنوز نرمال نباشد و CLT به‌خوبی صدق نکند. در این شرایط، نتایج آزمون‌های آماری قابل اعتماد نیستند.
2. همبستگی درون داده‌ها (مثلاً در نمونه‌های خوشه‌ای):  
   اگر مشاهدات در نمونه به‌طور مستقل انتخاب نشده باشند (مثلاً همه از یک کلاس یا منطقه جغرافیایی خاص باشند)، فرض استقلال که در CLT نقش دارد نقض می‌شود و خطای استاندارد محاسبه‌شده کمتر از مقدار واقعی خواهد بود.

در نتیجه قضیه حد مرکزی ابزار قدرتمندی برای استنتاج آماری است، اما فقط زمانی قابل اعتماد است که نمونه‌گیری به‌درستی و به‌شکل تصادفی و نماینده انجام شده باشد. در غیر این صورت، حتی با داده‌های زیاد هم ممکن است به نتایج نادرست برسیم.