

کارگاه روش‌های نمونه‌گیری و حجم نمونه

معاونت تحقیقات و فناوری - دانشگاه علوم پزشکی هرمزگان

دکتر شهرام زارع



بخش دوم

در پایان این کارگاه باید بتوانیم:

- استنباط صحیحی از نمونه داشته باشیم
- با روشهای مختلف نمونه گیری آشنایی نسبی پیدا کنیم
- استنباط صحیحی از اهمیت حجم نمونه داشته باشیم
- با عوامل مرتبط با حجم نمونه آشنا شویم
- با روشهای ساده محاسبه حجم نمونه آشنا شویم

برآورد حجم نمونه

۴

همواره یکی از مشکلات پژوهشگران در ابتدای هر پژوهش این است که از جامعه تحت مطالعه چند نمونه انتخاب کنیم تا نتایج بدست آمده را بتوان با اطمینان بالایی به جامعه آماری نسبت داد. از طرفی حجم نمونه مناسب از نظر تامین دقت نتایج و صرفه جویی در وقت و هزینه نیز اهمیت دارد. بزرگ بودن نمونه باعث هزینه و وقت زیاد و کوچک بودن آن موجب عدم دقت کافی در برآوردها می شود. روشهای متفاوتی بر اساس ملاحظات آماری برای برآورد حجم نمونه وجود دارد.

در تعیین حجم عوامل متعددی مؤثرند. از جمله:

- **واریانس:** اگر واریانس یا پراکندگی صفتی که می خواهیم آن را به وسیله نمونه گیری برآورد کنیم زیاد باشد، طبیعی است که به حجم نمونه بیشتری نیاز خواهیم داشت.
- **سطح اطمینان:** منظور انتظاری است که از نمونه انتخابی می رود که بتوان پارامترها را برآورد کرد و این انتظار را سطح اطمینان می نامیم و افزایش آن، افزایش حجم نمونه را در پی خواهد داشت.
- **با چه دقتی:** اینک با چه دقتی قرار است پارامترهای جمعیت هدف را برآورد کنیم، موضوعی است که در تعیین حجم نمونه مؤثر است. در واقع میزان خطای مورد پذیرش برای برآورد پارامتر جامعه آماری چقدر است؟
- **حجم جامعه آماری:** از آنجایی که باید صفت مورد بررسی به جامعه تعمیم داده شود، در یک جمعیت با حجم زیاد باید تعداد نمونه ها به قدری بزرگ باشند که ویژگی های جمعیت در آنها مشاهده شود و از طرفی تناسبی معقول بین تعداد کل جامعه آماری و حجم نمونه وجود داشته باشد.
- **توزیع آماری داده ها:** یکی دیگر از عوامل مؤثر بر تعداد حجم نمونه توزیع صفتی است که در جامعه وجود دارد و ما قصد برآورد آنرا داریم و اینکه چگونه می توانیم به آن جامعه دسترسی داشته باشیم.
- **و نهایتاً میزان اثر (Effect Size) و یا در مطالعات پزشکی تأثیر مهم از دیدگاه بالینی:** حداقل میزان تأثیر چقدر باشد تا انجام پژوهش از نظر بالینی اهمیت داشته باشد؟

چرا حجم نمونه؟

➤ تعیین حجم نمونه ممکن است به روش‌هایی که در ادامه به آن اشاره می‌شود، انجام پذیرد:

➤ **تعیین حجم نمونه براساس تجربه:** در این حالت محقق براساس اطلاعاتی که از توزیع احتمالی یا

نوع داده‌ها دارد، حجم نمونه را تعیین می‌کند. برای مثال در این حالت اگر حجم نمونه بیشتر از

حدّ مورد نیاز باشد، ممکن است فاصله اطمینان (Confidence Interval) ایجاد شده، دارای

دامنه ای بزرگتر از حد قابل قبول باشد که دقت برآورد را کاهش می‌دهد. همچنین با انتخاب حجم

نمونه بزرگتر از مقدار مورد نیاز، هزینه‌های تحلیل‌های آماری بدون آنکه ضرورتی داشته باشد

افزایش می‌دهد.

چرا حجم نمونه؟

- **تعیین حجم نمونه براساس میزان پراکندگی:** با انتخاب مقدار واریانس به عنوان معیار پراکندگی برای برآوردگر، می‌توان حجم نمونه را انتخاب کرد. در چنین حالتی، برای رسیدن به واریانس کوچک‌تر (خطای کمتر) برای برآوردگر، احتیاج به حجم نمونه بیشتری نیز هست. در نتیجه اگر هدف تعیین کران‌های فاصله اطمینان باشد، می‌توان با انتخاب حجم نمونه بزرگ، به طول فاصله اطمینان کوچک‌تری نیز دست یافت.
- **تعیین سطح معنی دار بودن:** با ثابت در نظر گرفتن میزان خطا، با افزایش مقدار سطح اطمینان یا سطح معنی دار بودن، به نمونه بیشتری نیز نیاز داریم. در نتیجه باید بین میزان خطا و سطح معنی‌داری به یک تعادل رسید تا نمونه مناسب بدست آید. دقت داشته باشید که با انتخاب میزان خطای بالا (دقت کمتر) و در نظر گرفتن سطح اطمینان مناسب، به حجم نمونه کمتری خواهیم رسید.
- در مسائل مربوط به تعیین حجم نمونه، اغلب توزیع داده‌ها را نرمال فرض می‌کنند. از طرفی می‌دانیم طبق قضیه حد مرکزی (Central Limit Theorem) و قانون اعداد بزرگ (law of Large Number) با افزایش حجم نمونه، توزیع‌های آماری (بشرط واریانس محدود) به سمت توزیع نرمال میل می‌کنند. بنابراین انتخاب توزیع نرمال برای داده‌ها کمی کاری غیر منطقی محسوب نمی‌شود بلکه فقط ممکن است حجم بزرگتری از نمونه به کار آید که باعث افزایش هزینه نمونه‌گیری می‌شود ولی در عمل خطای برآورد تغییری نخواهد کرد.

اما حجم نمونه چقدر باشد؟

➡ اما با همه این مسائل چنانچه هیچ محدودیتی از نظر زمان و یا هزینه انجام پژوهش و نمونه گیری نداشته باشیم، آیا بخصوص در مطالعات تحلیلی، مجاز به انتخاب تعداد نمونه به هر میزان دلخواه هستیم؟

➡ برای پاسخ به این سؤال بهتر است به نگاهی به فرایند انجام آزمونهای تصمیم گیری در آزمونهای آماری داشته باشیم.

چرا حجم نمونه (بخصوص در مطالعات تحلیلی) مهم هست؟

نتیجه گیری نهایی در همه آزمونهای آماری بر اساس مقدار عددی شاخص آماری آزمون که با استفاده از داده های جمع آوری شده از نمونه انتخابی محاسبه شده است و خطای نوع اول یعنی α و با توجه به توزیع آماری شاخص آماری آزمون انجام میگیرد که چنانچه این مقدار در ناحیه رد فرضیه H_0 (Rejection Region) یا ناحیه بحرانی قرار گیرد، آزمون معنی دار خواهد بود. عبارت دیگر فرضیه صفر رد و فرضیه جانشین یا فرضیه محقق تأیید می شود. در همه آزمونهای یک دنباله یا دو دنباله بهمین صورت است.

بعنوان مثال اگر به شاخص آماری آزمون در آزمون میانگین توجه کنید، این شاخص بصورت زیر است:

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

چرا حجم نمونه مهم است؟

➡ با توجه به مفروضات مسئله، به عنوان مثال برای نمونه $n=21$ ، این شاخص دارای توزیع t با $df = 20$ برای آزمون فرضیه (a یک مقدار ثابت)

$$H_0: \mu \leq a$$

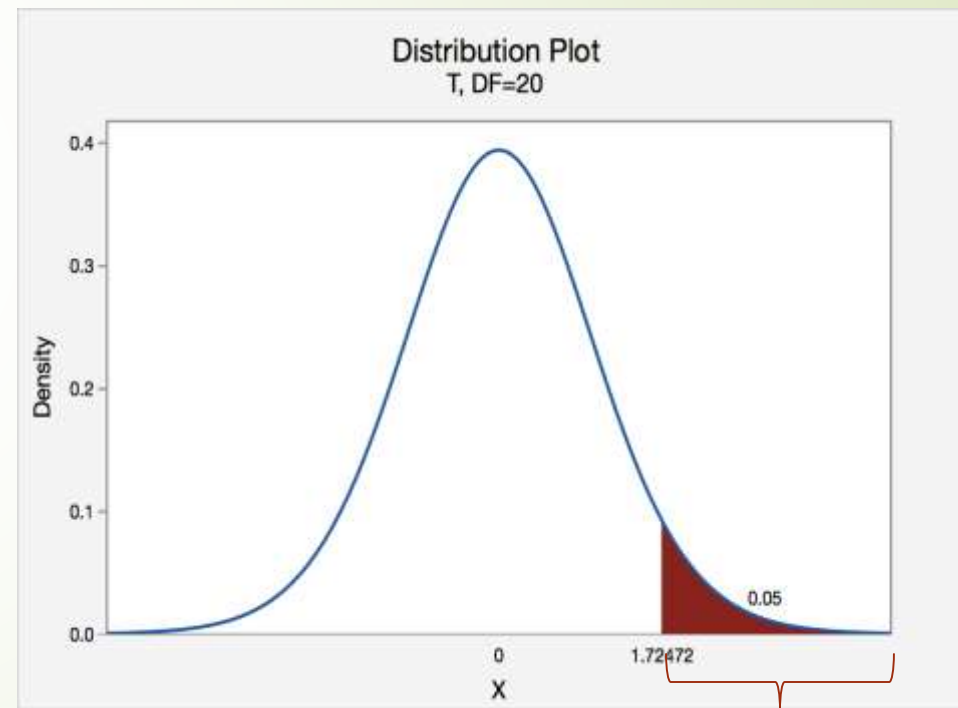
$$H_1: \mu > a$$

چنانچه مقدار شاخص آماری آزمون در ناحیه بحرانی قرار گیرد، آزمون معنی دار خواهد بود یعنی فرضیه صفر ردّ و فرضیه محقق پذیرفته میشود.

که در این مثال اگر مقدار شاخص آماری آزمون بزرگتر از مقدار

باشد (این مقدار از جدول توزیع t در سطح خطای 0.05 و درجه

۲۰ استخراج شده است)، آزمون معنی دار خواهد بود. حال اگر به شاخص آماری مربوط به این آزمون دقت کنید، مشاهده میشود که در



ناحیه بحرانی

چرا حجم نمونه مهم است؟

این فرمول با افزایش مقدار n (تعداد نمونه)، مقدار این شاخص نیز افزایش می یابد.

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

تعداد نمونه 

به این معنی که هر چقدر تعداد نمونه افزایش یابد، مقدار عددی این شاخص نیز افزایش می یابد. یعنی با افزایش تعداد نمونه احتمال معنی دار شدن آزمون نیز افزایش می یابد. بنابراین افزایش بی رویه تعداد نمونه ممکن است محقق را به نتایج غیر واقعی سوق دهد. بعبارت دیگر رد فرض صفر یا پذیرش فرضیه محقق، در حالی که این فرضیه در واقع صحیح نیست! (افزایش خطای نوع اول)

این موضوع در همه آزمونهای آماری بنوعی صادق است.

علاوه بر این، افزایش بی رویه تعداد نمونه می تواند محقق را از موضوع اصلی تحقیق منحرف سازد. یعنی با افزایش تعداد نمونه، موضوعات و تفاوت های کوچکتر خود را نمایان کنند طوری که محقق از موضوع اصلی پژوهش باز بماند.

■ بطور کلی دو رویکرد متفاوت در رابطه با محاسبه اندازه نمونه وجود دارد:

- رویکرد مبتنی بر دقت برآورد پارامترهای جمعیت Precision - based

- رویکرد مبتنی بر توان Power - based

- رویکرد مبتنی بر دقت:

فرض کنید هدف برآورد یک پارامتر نامعلوم جمعیت با دقت مشخصی باشد. در واقع باید فاصله اطمینان با دامنه مشخصی برای پارامتر مورد نظر داشته باشیم که بطور کلی این فاصله اطمینان با درجه مشخصی از اطمینان بشرح زیر خواهد بود:

$$\text{Estimate} \pm Z_{1-\frac{\alpha}{2}} \times SE \longrightarrow d$$

که در این فرمول: SE خطای معیار پارامتر مورد نظر و $Z_{1-\frac{\alpha}{2}}$ ضریب اطمینان خواهد بود. نکته قابل توجه این است که فرمول مربوط به SE همیشه شامل تعداد نمونه، n ، می باشد.

در فرمول فوق جزء $(Z_{1-\frac{\alpha}{2}} \times SE)$ تعیین کننده دامنه فاصله اطمینان است که کوچک یا بزرگ بودن آن که تعیین کننده میزان دقت برآورد است. اگر این مقدار را برابر با d (میزان دقت) در نظر بگیریم، یعنی

$$d = Z_{1-\frac{\alpha}{2}} \times SE$$

این مقدار حداکثر فاصله بین برآورد نقطه ای با کرانه پایینی و کرانه بالایی فاصله اطمینان است

با حل این معادله بسادگی می توان تعداد نمونه مورد نیاز را محاسبه کرد.

همانطور که مشاهده می شود در معادله فوق مقادیر SE و $Z_{1-\frac{\alpha}{2}}$ رابطه مستقیم با میزان دقت یعنی d دارد.

از سویی SE رابطه مستقیم با انحراف معیار یا واریانس متغیر مورد نظر دارد. بنابراین با افزایش این مقادیر یعنی با افزایش انحراف معیار و / یا با افزایش ضریب اطمینان، تعداد نمونه نیز افزایش پیدا می کند.

رویکرد مبتنی بر توان:

رویکرد محاسبه تعداد نمونه مبتنی بر توان آزمون، به آزمون فرضیه وابسته است که با توجه به خطای نوع اول (α)، خطای نوع دوم (β) و توان آزمون ($1-\beta$) محاسبه می گردد.

بر اساس این رویکرد در صورت وجود تفاوت (عدم صحت فرض صفر) با چه قدرت و دقتی می توانیم این تفاوت را به اثبات برسانیم؟ یعنی در واقع قدرت یافتن ارتباط و یا اختلاف واقعی چه میزان است؟

حجم نمونه بر اساس توان آزمون بصورت کلی برابر است با:

$$n = f(\alpha, \beta) \times \frac{\sigma^2}{d^2}$$

که $f(\alpha, \beta)$ برای مقادیر مشخصی از خطای نوع اول و خطای نوع دوم برابر است با:

α	β			
	0.05	0.1	0.2	0.5
0.05	13.0	10.5	7.9	3.8
0.01	17.8	14.9	11.7	6.6

برآورد حجم نمونه

سطح اطمینان و دقت توسط محقق تعیین می شود اما به طور معمول واریانس داده ها از قبل مشخص نیست. می توان با استفاده از روش های زیر واریانس را تخمین زد:

- از اطلاعات به دست آمده در مطالعات مشابه قبلی استفاده کرد.
- بر اساس یک نمونه گیری اولیه (پیش آزمون) واریانس را برآورد کرد.

گاهی در پرسشنامه ها که پاسخ ها کمی هستند و از کمترین و بیشترین مقادیر آنها اطلاع داریم، می توان انحراف معیار را از رابطه زیر محاسبه کرد:

$$\sigma = \frac{Max - Min}{6}$$

به عنوان مثال:

در مورد طیف پنج گزینه ای لیکرت، اگر ماهیت متغیر به گونه ای باشد که بخواهیم پاسخ ها را کمی در نظر بگیریم، انحراف معیار را به صورت زیر به دست می آوریم:

$$S = \frac{5-1}{6} = 0.667$$

توجه داشته باشید که از علامت σ برای انحراف معیار جمعیت و از علامت S برای انحراف معیار نمونه استفاده می شود که در واقع S برآورد یا تخمینی از σ می باشد.

برآورد حجم نمونه

برای برآورد تعداد نمونه ابتدا باید واریانس (σ^2) را با یکی از راههایی که قبلاً توضیح داده شد، به دست آوریم. در گام بعدی باید دقت برآورد را تعیین کنیم یعنی تعیین کنیم تا چه میزان در اندازه گیری صفت مورد نظر خطا مجاز است (d). سپس میزان اطمینان را (که اغلب اوقات ۹۵٪ در نظر گرفته می شود) مشخص کنیم و در نهایت از رابطه زیر تعداد نمونه لازم را محاسبه کنیم:

$$n = \frac{(z_{1-\frac{\alpha}{2}})^2 \times \sigma^2}{d^2}$$

در این رابطه، $z_{1-\frac{\alpha}{2}} = 1.96$

از جدول توزیع نرمال استاندارد و با اطمینان ۹۵٪ بدست می آید. برای میزان اطمینان های دیگر نیز میتوان این مقدار را از جدول بدست آورد. مثلاً برای اطمینان ۹۰٪ یا ۹۹٪.

برآورد حجم نمونه

مثال:

فرض کنید مدیر بهداشتی یک استان مایل است بداند میانگین زمان تماشای برنامه های بهداشتی تلویزیون در بین افراد ساکن یک منطقه چقدر است؟ برای این بررسی او به چه تعداد نمونه نیاز دارد تا با اطمینان ۹۵ درصد برآورد درستی از میانگین انجام داده باشد. او مقدار خطای اندازه گیری را ۵ دقیقه در نظر گرفته است و از مطالعات قبل می داند انحراف معیار زمان تماشای تلویزیون ۴۰ دقیقه است.

توجه داشته باشید که واریانس، مربع انحراف معیار می باشد.

با توجه به اطلاعات داده شده و استفاده از فرمول صفحه قبل، حد اقل تعداد نمونه مورد نیاز برابر است با:

$$n = \frac{(1/96)^2 \times (40)^2}{5^2} = 245/9 \approx 246$$

اصلاح جمعیت محدود

■ در مواقعی که با جمعیت محدود سروکار داشته باشیم میتوانیم با استفاده از ضریب اصلاح جمعیت محدود نسبت به تعدیل تعداد نمونه اقدام کرد. ضریب اصلاح جمعیت محدود بصورت زیر است:

$$n = \frac{n_0}{1 + \frac{n_0 - 1}{N}}$$

در این فرمول N تعداد جمعیت محدود و n_0 تعداد نمونه اولیه می باشد.

برآورد حجم نمونه

در حالتی که تعداد جمعیت هدف محدود است، برای تعیین حجم نمونه می توان از ضریب اصلاح جمعیت محدود تعداد نمونه را تعدیل کرد که نهایتاً فرمول محاسبه تعداد نمونه به صورت زیر خواهد بود.

پس از مشخص کردن واریانس، دقت برآورد و میزان اطمینان برای برآورد تعداد نمونه از رابطه زیر استفاده کنید:

$$n = \frac{N (z_{1-\frac{\alpha}{2}})^2 \times \sigma^2}{(N - 1)d^2 + (z_{1-\frac{\alpha}{2}})^2 \times \sigma^2}$$

برای مثال:

یک محقق برای بررسی وضعیت چرخه دانش در یک دانشگاه با ۲۴۰ عضو هیئت علمی می‌خواهد تعداد نمونه لازم را مشخص نماید. او در یک پیش آزمون انحراف معیار امتیاز مربوطه را ۱/۲ به دست آورده‌است و می‌خواهد این تحقیق را با اطمینان ۹۵٪ به سرانجام برساند. تعداد نمونه را چقدر در نظر گیرد وقتی مایل است خطای اندازه گیری در حد ۰/۰۳ باشد. به عبارت دیگر نتیجه بدست آمده حداکثر ۰/۰۳ واحد با مقدار واقعی در جمعیت اصلی تفاوت داشته باشد یا دقت اندازه گیری ما ۰/۰۳ باشد.

$$n = \frac{240 \times (1/96)^2 \times (1/2)^2}{239 \times (0/03)^2 + (1/96)^2 \times (1/2)^2} = 90$$

برآورد حجم نمونه

در صورتی که هدف برآورد میزان شیوع یا درصد یا نسبت یک ویژگی در یک جامعه باشد میتوان از فرمول زیر استفاده کرد. در این فرمول p برآورد مقدار آن ویژگی در جامعه است که مقداری بین صفر و یک است و با توجه به مطالعات گذشته تخمین زده می شود و $q = p - 1$ می باشد.

پس از مشخص کردن واریانس، دقت برآورد و میزان اطمینان برای برآورد تعداد نمونه از رابطه زیر استفاده می شود:

و اگر برآورد p مقدور نباشد، مقدار آنرا 0.5 در نظر می گیریم. زیرا بیشترین مقدار واریانس برای نسبت وقتی است که $p = 0.5$ باشد.

$$n = \frac{(z_{1-\frac{\alpha}{2}})^2 \times pq}{d^2}$$

برآورد حجم نمونه

محقق می خواهد با اطمینان ۹۵ درصد نسبت استفاده کنندگان از خدمات بهداشتی را در یک شهرستان برآورد کند. او می داند این نسبت در سال گذشته و یا بر اساس مطالعات مشابه ۲۱ درصد بوده است. محقق به چه تعداد نمونه نیاز خواهد داشت تا در تحقیق او، خطای اندازه گیری بیشتر از ۵ درصد نباشد.

$$\frac{1.96^2 * 0.21 (1 - 0.21)}{0.05^2} = 256$$

به عبارت دیگر باید حداقل ۲۵۶ نمونه از جمعیت هدف بطور تصادفی انتخاب شود.

برآورد حجم نمونه

زمانی که متغیر مطالعه دارای مقیاس رتبه ای دو ارزشی هستند مانند بلی و خیر یا موفقیت و شکست و حجم جامعه (N) نسبتاً کم است پس از مشخص کردن واریانس، دقت برآورد و میزان اطمینان برای برآورد تعداد نمونه از رابطه زیر استفاده می کنیم:

$$n = \frac{N (z_{1-\frac{\alpha}{2}})^2 \times pq}{(N - 1)d^2 + (z_{1-\frac{\alpha}{2}})^2 \times pq}$$

محقق می خواهد حجم نمونه را برای یک تحقیق پرسشنامه ای با طیف پنج گزینه ای لیکرت از جامعه ای با تعداد ۱۵۰ عضو، برآورد نماید. اگر سطح اطمینان در این تحقیق ۹۵٪ باشد و از نسبت موفقیت (p) اطلاعی در دست نباشد، حجم نمونه لازم را چه مقدار اختیار کند تا خطای اندازه گیری حداکثر ۰/۰۶٪ باشد؟

$$n = \frac{150 * 1.96^2 * 0.5^2}{(150 - 1) * 0.06^2 + 1.96^2 * 0.5^2} = 97$$

یک روش ساده برای تخمین واریانس در محاسبه حجم نمونه

همانگونه که شرح داده شد، برآورد انحراف معیار و یا تخمین احتمالی نسبت مورد مطالعه برای برآورد حجم نمونه در بیشتر موارد کار ساده ای نیست. برای آسانتر شدن کار میتوان میزان اثر یا (Effect Size) را کسری از انحراف معیار و یا کسری از فراوانی متغیر مورد نظر دانست و به این شکل صورت و مخرج کسر مربوط به محاسبه حجم نمونه را آسان ساخت. به عنوان مثال فکر کنید که میخواهیم میانگین وزن نوزادان را برآورد کنیم ولی هیچ تخمینی از انحراف معیار نداریم. برای استفاده از این روش میانبر فرض میکنیم که مثلاً d یکدهم مقدار انحراف معیار باشد. بنابراین فرمول محاسبه حجم نمونه بصورت زیر خواهد بود:

$$n = \frac{Z^2 \times SD^2}{(0.1 SD)^2} = \frac{Z^2}{(0.1)^2} = Z^2 \times 100$$

به همین ترتیب برای برآورد نسبت نیز می توان از این روش استفاده کرد

می توان فرض کرد مثلا مقدار d ، 0.2 برابر نسبت مورد نظر است که در این صورت فرمول مربوطه بصورت زیر خواهد شد:

$$n = \frac{z^2 * p(1-p)}{(0.2p)^2} = \frac{z^2 * (1-p)}{(0.04p)} = 25z^2 * \frac{1-p}{p} = 25z^2 * \left(\frac{1}{p} - 1\right)$$

و به این ترتیب برآوردی از حجم نمونه را مجاسبه کرد.

➤ اگر خوب به این دو فرمول جدید نگاه نمایید، متوجه خواهید شد در زمانی که مقدار انحراف معیار کوچک باشد و یا شیوع متغیر مورد نظر کم باشد و بخواهیم d ، کسری از انحراف معیار و یا شیوع باشد، حجم نمونه افزایش می یابد.

➤ به این نکته باید دقت داشت که مثلاً وقتی که بخواهیم یک متغیر با شیوع ۳۰٪ را برآورد نماییم، میزان d برابر ۳٪ یعنی برآورد دامنه اطمینان $3 \pm 30\%$ چندان غیردقیق نیست. ولی اگر میزان شیوع ۳٪ باشد، آنوقت میزان d برابر ۳٪ یعنی دامنه اطمینان $3 \pm 3\%$ زیاد دقیق به نظر نمی رسد و در این صورت برای دقیق عمل کردن باید مقدار d را بسیار کوچکتر فرض نمود. طبیعی است که در صورت ثابت فرض کردن مقدار d ، یعنی ۳٪، هرچه مقدار p به عدد ۵۰٪ نزدیک شود مقدار حجم نمونه افزایش می یابد. ولی اگر مقدار d را کسری از p بگیریم، با کوچکتر شدن مقدار p و به طبع کوچکتر شدن d ، حجم نمونه افزایش پیدا می کند.

محاسبه حجم نمونه برای مقایسه دو میانگین

با توجه به آنچه که قبلاً گفته شد چنانچه محاسبه اندازه نمونه را با رویکرد توان آزمون نیز محاسبه کنیم، فرمول مربوطه بصورت زیر خواهد بود:

$$n = \frac{2 (Z_{(1-\frac{\alpha}{2})} + Z_{(1-\beta)})^2 \sigma^2}{d^2}$$

مثلاً اگر بخواهیم تفاوت بین غلظت بیلروبین نوزادان در دو گروه را با حداکثر خطاهای نوع اول و دوم ۵٪ و ۱۰٪ بسنجیم و تخمین ما از انحراف معیار بیلروبین در خون نوزادان حدود ۴ باشد و به دنبال کشف حداقل ۲ واحد تفاوت (d) باشیم، حجم نمونه حدوداً ۸۵ نوزاد در هر گروه خواهد بود. در این مثال حجم نمونه کم است و دلیل آن زیاد بودن نسبت بین میزان اثر (۲) و انحراف معیار (۴) است. در توضیحات فوق بیان شد که معمولاً این نسبت حداکثر ۰/۲۵ در نظر گرفته میشود، ولی در حالت فوق این نسبت ۰/۵ است

محاسبه حجم نمونه برای مقایسه دو نسبت

زمانی که می‌خواهیم حجم نمونه را تعیین نماییم باید بیان کنیم که حداکثر خطای نوع اول و دوم مورد قبول چه میزان است. طبیعی است که هرچه بخواهیم دقیقتر کار نماییم، لازم است خطاهای نوع اول و دوم را کاهش دهیم و در نتیجه حجم نمونه بیشتری خواهیم داشت. ولی به صورت کلی معمولاً خطای نوع اول را ۵٪ و خطای نوع دوم را ۱۰٪ و در بعضی موارد ۲۰٪ در نظر می‌گیریم.

$$n = \frac{\left[Z_{\left(1-\frac{\alpha}{2}\right)} \sqrt{\bar{p}(1-\bar{p})} + Z_{(1-\beta)} \sqrt{p_0(1-p_0) + p_1(1-p_1)} \right]^2}{(p_1 - p_0)^2} \quad \bar{p} = \left(\frac{p_1 + p_0}{2} \right)$$

بر اساس این فرمول با توجه به خطای نوع اول و خطای نوع دوم، می‌توان حجم نمونه مورد نیاز را محاسبه کرد

تعیین حجم نمونه برای رگرسیون چندگانه

در تحلیل رگرسیون چندگانه با روش معمولی کمترین مجزورات استاندارد، باید به ازای هر متغیر پیش‌بین ۱۵ مورد در نظر گرفت.

با این حال در صورتی که شرایط اجازه می‌دهد، در صورتی که به ازای هر متغیر پیش‌بین تقریباً ۳۰ شرکت کننده وجود داشته باشد، محقق توان بهتری برای نمایان ساختن یک اندازه اثر کوچک خواهد داشت.

هنگامی که متغیر وابسته دارای اندازه اثر کوچک باشد، خطای اندازه‌گیری زیاد و یا از رگرسیون گام به گام استفاده شود، حجم نمونه باید بزرگتر انتخاب شود.

■ تعیین حجم نمونه برای تحلیل عاملی

- تحلیل عاملی یکی از رویکردهای آماری با نمونه بزرگ است که باید تعداد آزمودنی‌ها بیشتر از متغیرها باشد. نسبتهای مختلفی برای این منظور بیان شده است. یکی از آنها نسبت ۱۰ به ۱ است یعنی به ازای هر متغیر به ۱۰ نمونه نیاز داریم. در برخی منابع، در مورد تعیین حجم نمونه در تحلیل عاملی نسبت ۲۰ به ۱ یعنی حداقل ۲۰ نمونه به ازای هر متغیر پیشنهاد شده است.
- اشکال عمده این دستورالعمل‌ها این است که در تحلیل عاملی اکتشافی (EFA)، درباره نسبت بهینه تعداد موارد به تعداد شاخص‌ها اتفاق نظر وجود ندارد. حجم نمونه به مدل عاملی جامعه (یا واقعی) بستگی دارد. بویژه اینکه هنگامی که هر عامل دارای حداقل ۳ تا ۴ شاخص و متوسط اشتراکات در بین شاخصها ۰٫۷ و بالاتر باشد، تعداد موارد کمتری لازم است. یعنی نسبت ۱۰ به ۱ کافی است.
- البته به یاد داشته باشید که در تحلیل عاملی حجم نمونه کمتر از ۱۰۰ غیرقابل دفاع است. برخی معتقدند که حداقل حجم نمونه در تحلیل عاملی ۲۰۰ نفر است.

■ تعیین حجم نمونه برای تحلیل عاملی

■ برای به دست آوردن حجم نمونه در تحلیل عاملی تأییدی (CFA) از این نسبتها استفاده نمی‌شود. حجم نمونه در تحلیل عاملی تأییدی به تعداد پارامترها در کل مدل اندازه‌گیری بستگی دارد. این پارامترها عبارتند از: ضرایب الگو، واریانس‌های خطا، کوواریانس‌های خطا (برای خطاهای همبسته)، واریانس‌ها و کوواریانس‌های عامل‌ها. بنابراین تحلیل عاملی تأییدی مستلزم تعداد برآوردهای بیشتری است و در نتیجه به حجم نمونه بزرگتری نیاز داریم تا نتایج از دقت و معقولیت لازم برخوردار باشند.

■ حجم نمونه در تحلیل عاملی تأییدی به نوع روش برآورد و ویژگیهای توزیع نمونه نیز بستگی دارد. اگر روش برآورد حداکثر درستنمایی (Maximum Likelihood) و توزیع نرمال باشد، به حجم نمونه کمتری نیاز است که در این صورت نسبت ۲۰ به ۱ توصیه می‌شود.

■ به عنوان مثال در یک مدل تحلیل عاملی تأییدی با ۱۰ پارامتر ممکن است حجم نمونه ۲۰۰ نفر کافی باشد. اما در صورتی که جهت برآورد مدل روشی به غیر از حداکثر درستنمایی به کار برده شود و توزیع متغیرها دارای انحراف جدی از توزیع نرمال باشند، باید حجم نمونه بسیار بزرگتر باشد.

حجم نمونه در مطالعات کیفی

■ نمونه‌گیری هدفمند که نمونه‌گیری غیراحتمالی، هدفدار یا کیفی نیز نامیده می‌شود به معنای انتخاب هدف دار واحدهای پژوهش برای کسب دانش یا اطلاعات است. این نوع نمونه‌گیری به دنبال ایجاد قوانین ثابت و تغییرناپذیر و یا تعمیم نتایج نیست بلکه سعی در شناخت بهتر هر پدیده در زمینه خاص دارد. سه نوع عمده نمونه‌گیری هدفمند شامل نمونه‌گیری **برای قابلیت مقایسه**، نمونه‌گیری **موارد خاص یا یگانه** و نمونه‌گیری **متوالی** هستند. استفاده از روش اشباع داده در پژوهش‌های کیفی به عنوان استاندارد طلایی پایان نمونه‌گیری در نظر گرفته می‌شود.

حجم نمونه در مطالعات کیفی

- نمونه گیری در پژوهش‌های کیفی و کمی تفاوت بسیار دارد زیرا هدف آن به جای تعمیم یافته ها، کسب درک عمیق از پدیده مورد بررسی است. انتخاب مشارکت کنندگان در پژوهش کیفی با هدف دستیابی به بیشترین اطلاعات در مورد پدیده مورد بررسی انجام می‌شود. در پژوهش کمی تأکید بسیار زیادی بر انتخاب تصادفی و شانس برابر انتخاب همه اعضای جامعه پژوهش می‌باشد اما در پژوهش کیفی نمونه پژوهش یا مشارکت کنندگان، انتخاب یا دعوت می‌شوند.
- نمونه گیری در پژوهش‌های کیفی معمولاً با استفاده از روش‌های غیراحتمالی به صورت هدفمند انجام شده و تا اشباع نظری ادامه پیدا می‌کند. بیشتر تحقیقات کیفی با استفاده از ابزار مصاحبه و از دیدگاه خبرگان انجام می‌شود. تعداد نمونه نیز معمولاً بسیار محدود است و ملاک پایان نمونه‌گیری اشباع نظری است.

حجم نمونه در مطالعات کیفی

- نمونه گیری در پژوهش‌های کیفی بسیار مهم است. اگر موضوع مورد مطالعه خیلی کلی و عمومی است برای نمونه گیری در پژوهش‌های کیفی از روش **نمونه گیری گلوله برفی** استفاده کنید. اگر موضوع مورد مطالعه در خصوص یک ویژگی خاص یا یک حوزه مشخص است برای نمونه‌گیری از روش **نمونه گیری هدفمند** استفاده کنید.
- در برخی از منابع مطالعات برای پژوهش‌های کیفی نمونه‌ای بین ۵ تا ۲۵ واحد در نظر گرفته اند. با این وجود دقت داشته باشید فرایند مصاحبه تا رسیدن به اشباع نظری ادامه پیدا می‌کند. اشباع نظری نقطه ای است که پژوهشگر درمی‌یابد دیگر حرف تازه‌ای زده نمی‌شود.

بطور خلاصه ...

- در پژوهشهای توصیفی، هر چه حجم نمونه بزرگتر باشد، طبیعتاً میزان اشتباهات در نتیجه‌گیری کاهش می‌یابد و بر عکس. حجم نمونه ارتباط بسیار نزدیکی با آزمون فرضیه صفر دارد. بدین ترتیب که هر چه اندازه گروه نمونه بزرگتر باشد محقق با قاطعیت بیشتری فرض صفر را که واقعاً نادرست است رد می‌کند.
- هر قدر حجم جامعه کوچکتر باشد نسبت بیشتری از جامعه باید در نمونه وجود داشته باشد و هر قدر حجم جامعه بزرگتر باشد نسبت کمتری از جامعه باید در نمونه وجود داشته باشد. اگر حجم جامعه ۳۰ نفر یا کمتر باشد محقق تقریباً باید کل جامعه را انتخاب می‌کند.
- اگر حجم جامعه بزرگ باشد، باید نمونه بزرگتری انتخاب شود. همچنین توجه داشته باشید که با افزایش حجم جامعه، حجم نمونه با میزان کمتری افزایش می‌یابد. در حجم جامعه بالاتر از ۳۸۰ نفر، حجم نمونه تقریباً ثابت می‌ماند.
- هر چه جامعه ناهمگون‌تر و یا به عبارت دیگر واریانس آن بیشتر باشد، محقق باید نمونه بزرگتری را انتخاب کند.

■ باید توجه داشت که در تحقیقاتی که احتمال ریزش و افت آزمودنی‌ها وجود دارد، باید همیشه نمونه ای بزرگتر از آنچه واقعاً نیاز هست انتخاب کرد. افت آزمودنی‌ها به ویژه در تحقیقات پانل (panel) روی می دهد. تحقیقاتی که در آن یک گروه از آزمودنی‌ها در طول زمان چندین بار مورد اندازه گیری قرار می گیرند. معمولاً محقق باید قبل از انجام تحقیق انتظار ۱۰ تا ۲۵ درصد ریزش نمونه را داشته باشد.

■ حجم نمونه تا حد زیادی به هدف و روش تحقیق بستگی دارد. در تحقیقات کیفی معمولاً از نمونه کوچک استفاده می شود. برای پژوهش‌های توصیفی، نمونه‌ای به حجم حداقل ۱۰۰ نفر نیاز است. در پژوهش‌های همبستگی برخی منابع حداقل حجم نمونه را ۳۰ نفر و برخی دیگر ۵۰ نفر ذکر کرده‌اند. در پژوهش‌های از نوع آزمایشی و علی-مقایسه‌ای، حجم نمونه حداقل ۱۵ نفر در هر گروه توصیه می شود. در تحقیقاتی که نیاز به طبقه‌بندی جامعه برای نمونه‌گیری می باشد، حداقل نمونه هر طبقه بین ۲۰ تا ۵۰ نفر است.

■ در تحقیقات چندمتغیره، حجم نمونه باید چند برابر (ترجیحاً ۱۰ برابر) تعداد متغیرها در پژوهش باشند.



پیروز باشید!