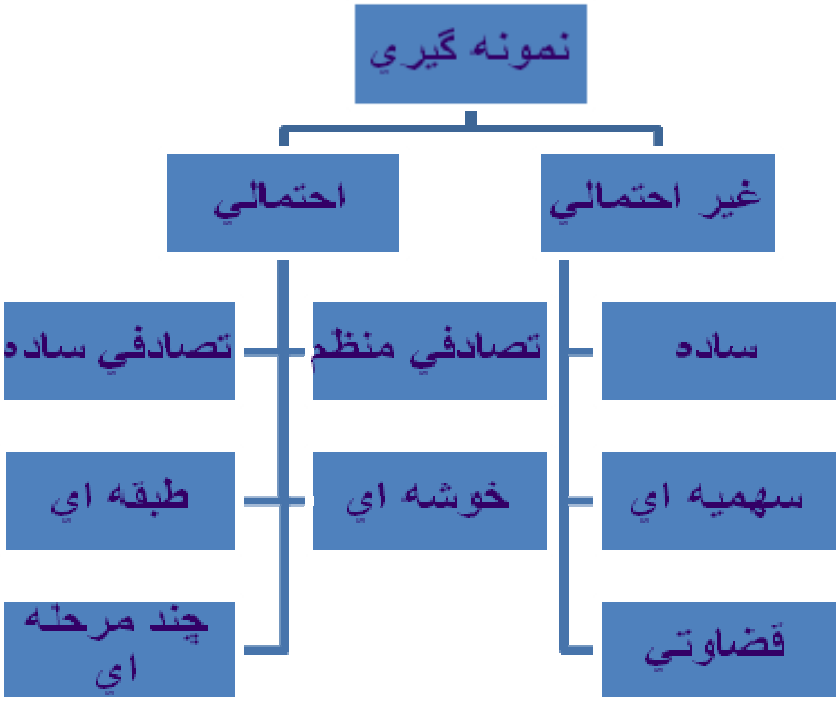
فصل اول

روش‌های نمونه‌گیری

مقدمه

شرط اساسی یک بررسی خوب انتخاب یک روش نمونه‌گیری صحیح است. نمونه انتخاب شده باید معرف جامعه باشد. اولین گام در طراحی یک نمونه‌گیری این است که واحد مطالعه را تعریف نماییم.



روش‌های نمونه‌گیری غیر احتمالی

در پژوهش‌های پزشکی و غیرپزشکی، بسیار اتفاق می‌افتد که پژوهشگر نمی‌تواند نمونه‌ای انتخاب کند که معرف جامعه باشد. موارد استفاده آن در مطالعات مقدماتی یا در پژوهش‌هایی است که برای بررسی مشکلات احتمالی یک پژوهش خاص در محیطی محدود می‌باشد که دو مورد از رایج‌ترین روش‌ها در زیر آورده شده است:

1. نمونه‌گیری آسان

در این نوع نمونه‌گیری، پژوهشگر از نمونه‌های در دسترس استفاده می‌کنند. مثل استفاده پزشک از بیماران مراجعه-کننده به مطب شخصی خود برای مقایسه اثر دو دارو در درمان یک بیماری

2. نمونه‌گیری سهمیه‌ای

در این نوع نمونه‌گیری هدف جامعیت بخشیدن به نمونه دردسترس و در صورت امکان رعایت تناسب موجود در جامعه می‌باشد. در واقع نمونه‌گیری سهمیه‌ای تلاشی برای بهبود شرایط نمونه‌گیری آسان است.

نکته:

استفاده از نمونه‌گیری‌های غیر احتمالی در مطالعه‌های مقطعی (مطالعه‌های شیوع)، که انتظار می‌رود مخرج کسر نماینده خوبی از جامعه باشد، اشتباه بزرگی است.

روش‌های نمونه‌گیری احتمالی

نمونه‌گیری تصادفی ساده (Simple random sampling)

- نمونه‌گیری تصادفی ساده از این جهت حائز اهمیت است که پایه ای فراهم می‌کند که نظریه آماری نمونه‌گیری بر پایه آن بنا شده است.
- در عمل همه نمونه گیری‌ها بدون جایگزینی انجام می‌شود.
- در این روش تمام اعضای جامعه دارای شانس مساوی یا حداقل شانس معینی برای قرار گرفتن در نمونه خواهند بود.
- در نمونه‌گیری تصادفی ساده شانس انتخاب هر نمونه n/N است.

روش انجام نمونه‌گیری:

- مشخص کردن جامعه هدف.
- تهیه لیست به روز شده برای انجام نمونه‌گیری.
- به تمام جامعه هدف یک شماره داده شود .
- با یکی از روش‌های احتمالی نمونه‌ها انتخاب شوند.

مزایای نمونه گیری تصادفی ساده :

- سادگی روش
- خطای نمونه‌گیری به سادگی قابل اندازه‌گیری است.
- برای نمونه‌گیری از جوامع کوچک روشی مناسب و عملی است.
- سوگیری انتخاب در این روش وجود ندارد.

معایب نمونه‌گیری تصادفی ساده :

- نیاز به لیست کامل و روز آمدی از جامعه دارد.
- نمونه ممکن است بسیار پراکنده باشند و دسترسی به همه نمونه‌ها امکان‌پذیر نباشد.
- هزینه‌بر است.
- به خوبی معرف جامعه دارای سطوح مختلف متغیر نیست.

چگونگی انتخاب عدد تصادفی :

- جدول اعداد تصادفی
- استفاده از برنامه ها کامپیوتری مثل EXCEL
- استفاده از کیسه و قرار دادن شماره ها در آن
- استفاده از روز تولد یا شماره سریال پول
- استفاده از کارت

استفاده از جدول اعداد تصادفی :

1 ـ تعداد ارقام حجم نمونه مشخص گردد.

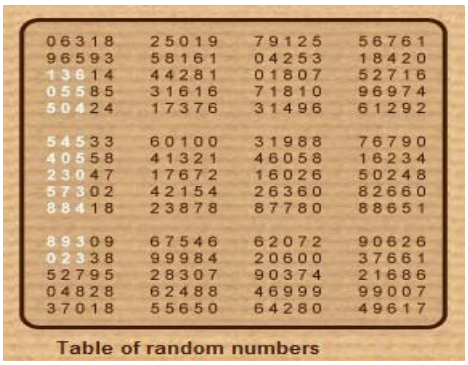
2 ـ بطور تصادفی (نوک خودکار و ...) نقطه‌ای را در روی جدول اعداد تصادفی پیدا کنید.

3 ـ جهت حرکت در جدول را مشخص کنید (مثلاً از راست به چپ).

4 ـ با توجه به تعداد ارقام حجم نمونه از نقطه تصادفی شروع به حرکت در جهت تعیین شده نمائید.

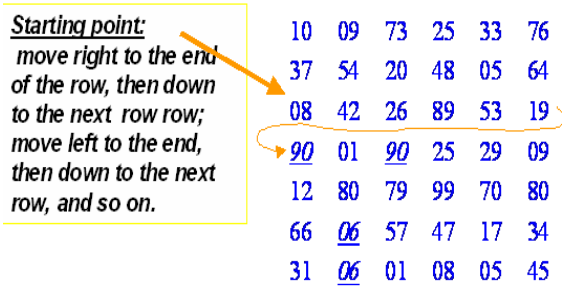
5 ـ تمام اعداد باید به اندازه رقم‌های حجم نمونه مورد نیاز خوانده شوند، مثلاً اگر حجم نمونه سه رقمی باشد عدد2 بصورت002خوانده می‌شود.

6 ـ اینکار را ادامه دهید تا به اندازه مورد نیاز انتخاب نمایید.



06318	25019	79125	56761
96593	58161	04253	18420
13614	44281	01807	52716
05585	31616	71810	96974
50424	17376	31496	61292
54533	60100	31988	76790
40558	41321	46058	16234
23047	17672	16026	50248
57302	42154	26360	82660
88418	23878	87780	88651
89309	67546	62072	90626
02338	99984	20600	37661
52795	28307	90374	21686
04828	62488	46999	99007
37018	55650	64280	49617

Table of random numbers



Starting point:
move right to the end of the row, then down to the next row row; move left to the end, then down to the next row, and so on.

10	09	73	25	33	76
37	54	20	48	05	64
08	42	26	89	53	19
90	01	90	25	29	09
12	80	79	99	70	80
66	06	57	47	17	34
31	06	01	08	05	45

نمونه‌گیری با روش منظم (Systematic random sampling)

- دراین روش تمام اعضای جامعه دارای شانس مساوی یا حداقل شانس معینی برای قرار گرفتن در نمونه خواهند بود.
- نمونه‌گیری با روش منظم چه به تنهایی و چه در ترکیب با برخی روش‌های دیگر می‌تواند متداولترین روش نمونه‌گیری باشد.

بسیار شبیه روش نمونه‌گیری ساده است با دو تفاوت :

1- در نمونه‌گیری با روش منظم تنها یک عدد تصادفی در کل پروسه نمونه‌گیری لازم است.

2- نیاز به شناخت همه افراد در جامعه نمونه نیست.

روش انجام نمونه‌گیری:

1- چهارچوب نمونه‌گیری تهیه می‌شود.

2- حجم نمونه تعیین می‌شود.

3- فاصله نمونه‌گیری تعیین می‌شود.

فاصله نمونه‌گیری = اندازه نمونه‌جمعیت مورد مطالعه/حجم نمونه

4- انتخاب اولین نمونه بر اساس روش تصادفی بین1 تا عدد کسر نمونه‌گیری صورت می‌گیرد.

5- نمونه‌های بعدی با اضافه نمودن کسر نمونه‌گیری به نمونه ماقبل بدست می‌آیند.

مثال:

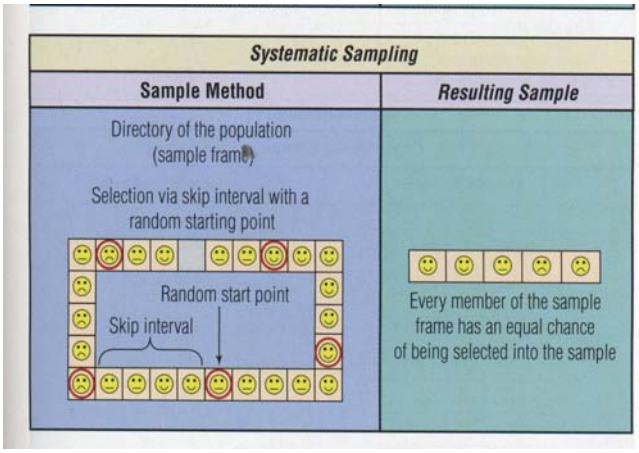
در یک جامعه 1000نفری قراراست50 نمونه جهت پایش انتخاب شوند:

کسر نمونه‌گیری = 1000 / 50 = 20

فرض کنید بطور تصادفی اولین نمونه عدد 8 انتخاب شده‌است.

نمونه دوم = 20+8=28

نمونه سوم = 28+20=48



مزایای نمونه‌گیری با روش منظم :

- سادگی روش
- نیاز به شناخت تک تک جامعه نمونه نیست.
- کم هزینه تر و عملی تر از نمونه‌گیری تصادفی ساده است .

معایب نمونه‌گیری با روش منظم :

- به خوبی معرف جامعه دارای سطوح مختلف متغیر نیست.
- اگر فاصله نمونه‌گیری بر تغییرات منظم در درون جامعه مورد پایش منطبق شود، تورش رخ می‌دهد.

نمونه‌گیری تصادفی طبقه‌بندی شده(Stratified random sampling)

با توجه به معایب نمونه‌گیری تصادفی ساده و منظم، در صورت امکان تقسیم جامعه مورد پژوهش به طبقه‌ها یا گروه‌ها می‌توان نمونه‌گیری را در هر طبقه انجام داد.

معایب نمونه‌گیری تصادفی ساده:

- نیاز به لیست کامل و روز آمدی از جامعه دارد.
- نمونه ممکن است بسیار پراکنده باشند و دسترسی به همه نمونه‌ها امکان پذیر نباشد.
- هزینه‌بر است.
- به خوبی معرف جامعه دارای سطوح مختلف متغیر نیست.

معایب نمونه‌گیری با روش منظم:

- به خوبی معرف جامعه دارای سطوح مختلف متغیر نیست.
- اگر فاصله نمونه‌گیری بر تغییرات منظم در درون جامعه مورد پایش منطبق شود، تورش رخ می‌دهد.

در نمونه‌گیری تصادفی ساده و منظم نمی‌توان اطمینان داشت که در نمونه مورد نظر، نسبت کسانی که دارای ویژگی های معین هستند برابر همان نسبت در جامعه باشند.

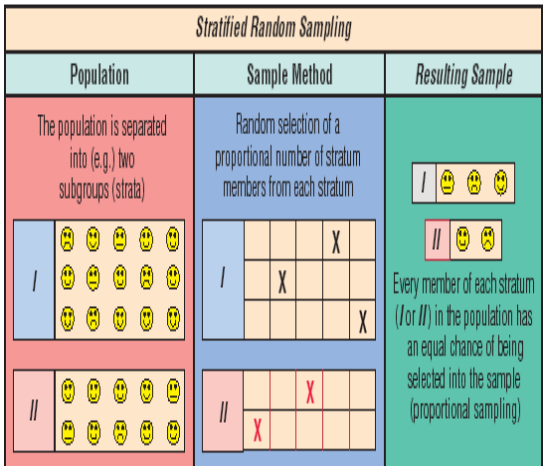
برای حل این مشکل، ممکن است با استفاده از اطلاعات کمکی بتوان جامعه را به طبقه‌ها یا گروه‌هایی تقسیم کرد در آن صورت می‌توان نمونه‌گیری را در هر طبقه انجام داد البته زمانی می‌توانیم از این نوع نمونه‌گیری استفاده کنیم که نسبت هر طبقه در جامعه مشخص باشد.

اگر بتوان جامعه مورد مطالعه را بر اساس صفت مشخصی طوری به گروه‌های جدا از هم (بدون همپوشانی) تقسیم کرد که تعداد افراد هر گروه معلوم باشد می‌توان نمونه‌گیری با طبقه‌بندی را با انجام نمونه‌گیری تصادفی ساده در هر طبقه و ادغام نتیجه در هم انجام داد.

در عین حال یکی از دلایل دیگر برای انتخاب این نوع نمونه گیری آن است که با استفاده از این نوع نمونه‌گیری می‌توان پارامترهای مربوط را برای هر طبقه برآورد کرد.

نمونه‌گیری تصادفی طبقه‌بندی شده

روشی از انتخاب نمونه است که در آن جمعیت را بر اساس یک سری ویژگی‌ها طبقه‌بندی می‌کنند که این طبقات هیچ نوع همپوشانی نداشته باشند یعنی هر فرد فقط و فقط به یکی از طبقات اختصاص داشته باشد سپس با استفاده از روش نمونه‌گیری اتفاقی ساده از هر کدام از این طبقات نمونه مورد نیاز را انتخاب می‌کنیم و در پایان نمونه‌ها را روی هم می‌ریزیم تا پارامتر جمعیت را برآورد کنیم.



برای بدست آوردن یک برآورد با دقت بالا، عناصر در درون طبقات باید تا حد ممکن گوناگونی کمی داشته باشند یا اصطلاحا هموژن باشند و گوناگونی بین طبقات بطور نسبی زیاد باشد .

با وجود توضیحات قبل نتیجه می‌گیریم که Stratification random sampling حالتی از نمونه‌گیری است که خصوصیات زیر را داشته باشد :

1- جمعیت تحت مطالعه دارای N عضو است.

2- جمعیت از لحاظ داشتن صفات خاصی به H طبقه متفاوت تقسیم می‌شوند.

3- هر عضو فقط می‌تواند به یکی از طبقات تعلق داشته باشد. (بدون همپوشانی)

4- تعداد مشاهدات در هر طبقه مشخص است و با N_h مشخص می‌شود.

5 - تعداد کل جمعیت برابر است با مجموع تعداد جمعیت‌ها در هر طبقه.

6- محقق با استفاده از به کارگیری یکی از روش‌های نمونه‌گیری مثل نمونه‌گیری اتفاقی ساده، از هر طبقه تعداد نمونه مورد نیاز را انتخاب می‌کند.

بعد از اینکه طبقات را مشخص کردیم برای انتخاب نمونه از هر طبقه 2راه وجود دارد:

1- انتخاب تعداد مساوی از هر طبقه

2- انتخاب نمونه از هر طبقه با در نظر گرفتن نسبت آن طبقه در جمعیت .

راه اول و ساده‌تر این است که تعداد مساوی از هر طبقه انتخاب کنیم $n_h = n/l$ که L برابر است با تعداد طبقات و n_h برابر است با تعداد عناصر انتخاب شده از طبقه h ام و n هم کل تعداد نمونه است.

مثال: یک جمعیت 1000 نفری را از نظر مقطع تحصیلی به 4 طبقه تقسیم کرده‌ایم که طبقات شامل 400، 300، 200 و 100 هستند و می خواهیم از بین آنها 100 نفر را به عنوان نمونه انتخاب کنیم تعداد افراد هر نمونه برابر می شود با $n_h = \frac{100}{4}$ و به این ترتیب از هر طبقه ما 25 نفر را انتخاب می کنیم.

$\frac{25}{100}$ ، $\frac{25}{200}$ ، $\frac{25}{300}$ ، $\frac{25}{400}$.

راه دوم: روش مفید تر انتخاب افراد در طبقات اختصاص تناسبی proportional Allocation است. در این روش کسر نمونه‌گیری $\frac{n_h}{N_h}$ مختص هر طبقه است و در مثال قبل برای طبقات برابر است با :

$$n_h = N_h\left(\frac{n}{N}\right)$$

- $n_1=(400)\times(\frac{100}{1000})= 40$
- $n_2=(300)\times(\frac{100}{1000})=30$
- $n_3=(200)\times(\frac{100}{1000})=20$
- $n_4=(100)\times(\frac{100}{1000})=10$

مثال: فرض کنید در یک دانشگاه 1000 دانشجو در سه رشته پزشکی، بهداشت و پرستاری مشغول تحصیل باشند و ما بخواهیم میانگین مجموع DMF یا نسبت عینکی بودن آنها را بر آورد کنیم. تعداد کل، تعداد نمونه، میانگین و نسبت حاصل از نمونه‌گیری برای هر گروه در جدول شماره 1 خلاصه شده است.

جدول شماره 1 : مشخصات و حاصل یک نمونه‌گیری طبقه ای:

رشته تحصيلي	نسبت افراد عينكي	ميانگين DMF	تعداد نمونه	تعداد كل
پزشكي	$\frac{1}{2}$	6	40	500
بهداشت	$\frac{1}{3}$	8	40	300
پرستاري	$\frac{1}{4}$	7	40	200

در این صورت با وجود مساوی بودن تعداد نمونه ، میانگین گرفتن از میانگین‌های DMF، یعنی $\frac{1}{3}(7+8+6) = 7$ برآورد بدون سوگرایی از میانگین جامعه را بدست نخواهد داد.

میانگین نسبت ها یعنی $\frac{1}{3}\left(\frac{1}{4}+\frac{1}{2}+\frac{1}{3}\right)=0.36$ نیز برآوردی بدون سوگیری برای نسبت عینکی بودن در جامعه 1000 نفری به دست نمی‌دهد.

نمونه‌گیری خوشه‌ای(Cluster sampling)

نمونه‌گیری خوشه‌ای یکی از انواع نمونه‌گیری دو مرحله‌ای است.

- برای انجام آن بعد از تقسیم بندی کل جمعیت به منطقه‌های جدا از هم
- چند منطقه بطور تصادفی انتخاب می‌شوند
- کلیه افراد موجود در منطقه‌های انتخاب شده شانس انتخاب شدن پیدا می‌کنند.

روش کار

1. ابتدا نقشه شهر یا محیط مورد پژوهش را بر حسب تقسیم بندی‌های سیاسی یا دیگر انواع تقسیم بندی‌ها (منطقه های وسیع، ناحیه‌های شهری، آموزشی، پستی و ...) تهیه و تنظیم می‌کنند.
2. تک تک منطقه‌های وسیع را به واحدهای کوچک‌تر و فرعی‌تر تقسیم می‌کنند.

واحدهای کوچک‌تر و فرعی‌تر = خوشه

توجه:

معمولاً تلاش می‌شود تا در هر منطقه وسیع، خوشه‌های مناسب مشخص شوند به طوری که عوامل و واحدهای درونی هر خوشه بطور فیزیکی یا منطقه‌ای به هم نزدیک باشند.

سوال:

چند خوشه از بین خوشه‌های شماره‌گذاری شده باید انتخاب کرد؟

بطور کلی می‌توان گفت اگر همه خوشه‌ها به صورت ایده آل ویژگی‌های نزدیک به هم، به‌ویژه در مورد متغیرهای کمی، واریانس کاملاً مشابه داشته باشند. یک خوشه به نمایندگی همه خوشه‌ها کافی خواهد بود و هر چه این مشابهت کمتر باشد تعداد خوشه‌های بیشتری لازم است.

چگونه نگرانی از نامشابه بودن خوشه‌ها را رفع کنیم؟

برای رفع این مشکل معمولاً تعداد نمونه لازم را در نمونه‌گیری خوشه ای 1/5 تا 2/5 برابر می‌کنیم.

فرض کنید: یک بررسی برای مطالعه مراقبت‌های پیش از زایمان زنان باردار در یک شهر بزرگ طراحی شده است.

مشکلات متعدد در چنین مطالعه‌ای بدین شرح هستند:

1. جامعه خیلی بزرگ است و احتمالاً تهیه یک چارچوب دقیق و درست غیرممکن است. حتی اگر کسی بتواند چنین کاری انجام دهد هزینه لازم جهت جمع‌آوری اطلاعات دقیق و به روز و با جزئیات از تک تک افراد جامعه بسیار گران خواهد بود.
2. جمعیت بسیار متفرق هستند (محدودیت در زمان موجود و هزینه برای سفر جهت جمع‌آوری اطلاعات).

یک راه حل برای چنین مشکلاتی استفاده از نمونه‌گیری خوشه‌ای است.

- تکنیک‌های دیگر نمونه‌گیری اتفاقی ساده و نمونه‌گیری منظم، به یک چارچوب نمونه‌گیری که در آن لیست افراد و واحدهای سرشماری شده مشخص هستند احتیاج دارند.

- بعضی اوقات، به‌ویژه در بررسی‌ها بر روی جمعیت‌های انسانی گردآوری چارچوب نمونه‌گیری از همه واحدهای سرشماری شده برای کل جامعه عملی نیست. از طرف دیگر اغلب چارچوب نمونه‌گیری فقط در حد گروه های شناسایی شده مشخص و یا خوشه‌هایی از واحدهای مورد مطالعه موجود هستند، بدون آنکه لیست روشن و دقیقی از تک تک افراد سرشماری شده وجود داشته باشد.

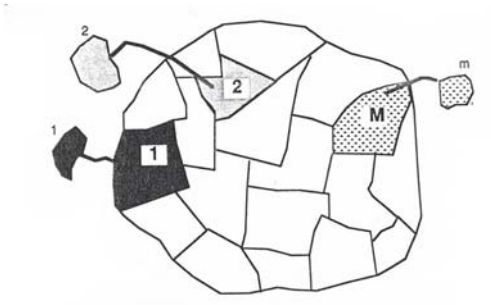
نمونه‌گیری می‌تواند از چنین چارچوب‌هایی نیز انجام پذیرد:

- یک نمونه از کل خوشه‌ها گرفته شود.
- یک لیست از واحدهای سرشماری شده فقط برای آن خوشه‌هایی که در نمونه انتخاب شده‌اند بدست آوریم.
- در نهایت یک نمونه از واحدهای سرشماری شده مشخص انتخاب کنیم.
- روندی که یک نمونه از واحدهای لیست شده انتخاب می‌شود قدم به قدم است.

برای مثال اگر خوشه‌ها بلوک‌های شهری هستند 2 قدم برای انتخاب نمونه از خانواده‌ها وجود دارد:

قدم اول) یک نمونه از بلوک‌ها انتخاب می‌کنیم.

قدم دوم) یک نمونه از خانواده‌های داخل هر بلوک انتخاب شده در قدم اول، انتخاب می‌کنیم.



Schematic representation of cluster sampling

توجه:

- وقتی از اصطلاح "خوشه" در متدلوژی بررسی نمونه استفاده می‌شود منظور هر واحد نمونه‌گیری است که شامل یک یا بیشتر از یک واحد لیست شده است.
- در این نوع نمونه‌گیری گروه‌هایی از عناصر(خوشه ها) بجای تک تک افراد به عنوان نمونه استفاده می‌شوند.

- در این نوع نمونه‌گیری از یک چارچوب شامل چندین خوشه از واحدهای جمعیتی استفاده کند.
- در اصطلاحات نمونه‌گیری به این قدم‌ها "مرحله" گویند و طرح‌های نمونه‌گیری اغلب به وسیله تعداد مراحلی که دارند طبقه‌بندی می‌شوند.
- نمونه‌گیری خوشه‌ای تک مرحله‌ای نوعی است که در آن نمونه‌گیری تنها در یک قدم انجام شده. یعنی یک مرتبه یک نمونه از خوشه‌ها انتخاب شده و تمام افرادی که در خوشه‌های انتخاب شده قرار داشتند به عنوان نمونه در نظر گرفته می‌شوند.

نمونه‌گیری چند مرحله‌ای (Multistage Sampling)

نمونه‌گیری دو مرحله‌ای می‌تواند به مراحل بیشتری گسترش یافته و هر یک از روش‌های نمونه‌گیری احتمالی را به کار بندد که در این صورت نمونه‌گیری چند مرحله‌ای (Sampling Multistage) نامیده می‌شود.

بطور کلی نمونه‌گیری مجدد از خوشه‌های انتخاب شده به دلایل زیر صورت می‌گیرد:

1. ممکن است خوشه‌های طبیعی وجود داشته باشد که نمونه‌گیری را آسان کند ولی این خوشه‌ها بزرگتر از اندازه اقتصادی مورد نظر باشد. مثال: محدوده‌های بلوک‌های شهری

2. می‌توان از هزینه ایجاد خوشه‌های کوچکتر در کل جمعیت اجتناب کرد و آن را به واحدهای نمونه انتخاب شده محدود کرد. مثال: ایجاد تقسیم‌بندی‌های مناطق کوچک برای کل یک شهر یا کشور
3. معمولا هر چه خوشه بزرگتر باشد واریانس آن کوچکتر است. مثال: انتخاب یک خوشه شامل 4 خانه کنار هم در یک محل
4. نمونه‌گیری از خوشه‌های متراکم ممکن است مشکلات عملی ایجاد کند. مثال: مصاحبه مجزای تمام اعضای یک خانواده

مثال:

1. برای بررسی کشاورزان یک استان می‌توان با نمونه‌گیری از شهرستان‌های استان شروع کرد و سپس از دهستان-های هر شهرستان منتخب نمونه‌گیری کرد. در داخل هر دهستان می‌توان فهرستی از آبادی‌ها تهیه و از آن فهرست نمونه‌گیری کرد، برای هریک از آبادی‌های انتخاب شده نیز می‌توان فهرستی از کشاورزان ساکن بدست آورده و نمونه‌ای از آنها برگزید. در این روش کشاورزان نمونه نهایی ما هستند که انتخاب آن‌ها متضمن انتخاب چند نمونه مختلف است و به گونه‌ای انجام می‌شود که هزینه نهایی بررسی به حداقل برسد. نتیجه این روش نمونه‌گیری تمرکز مصاحبه در چند نقطه جغرافیایی و کاهش هزینه رفت و آمد است.

اصل کلی، به حداکثر رساندن خوشه‌های اولیه منتخب و متعاقباً انتخاب تعداد نسبتاً معدودی افراد یا واحد ازهرخوشه منتخب است.

2. بررسی تأثیرپذیری اجرای طرح توانبخشی مبتنی بر جامعه به نحوه نگرش خانواده‌های دارای معلول روستایی استان اردبیل

- جامعه آماری شامل کلیه خانوارهای روستایی استان اردبیل است.
- واحد نمونه‌گیری پژوهش «خانوار روستایی دارای معلول» می‌باشد.

در مرحله اول از بین شهرستان‌های استان که شامل 9 شهرستان می‌شود به تصادف 5 شهرستان گزینش شدند (شهرستان‌های خلخال، نمین، نیر، مشگین‌شهر و گرمی).

در مرحله دوم از بین شهرستان‌های گزینش شده به تناسب جمعیت آنها چند بخش به تصادف انتخاب شدند.

در مرحله سوم از بین بخش‌ها روستاهای مورد نظر به تصادف گزینش شدند.

در مرحله چهارم در روستاهای گزینش شده از میان خانوارهای روستایی دارای معلول سه خانوار از هر روستا بصورت تصادفی انتخاب شدند.

در مرحله پنجم از هر خانوار دو نفر به تصادف مورد ارزیابی قرار گرفتند.

فصل دوم

تعیین حجم نمونه

مقدمه

برآورد حجم نمونه گامی اساسی و مهم در انجام یک پژوهش است کـه اکثـر محققـین در شـروع مطالعـه خود با آن سروکار دارند. چنانچه حجـم نمونـه کمتـر از میـزان لازم در نظـر گرفته شود، ممکن است نتـایج اسـتنباط شـده از آن در مـورد جامعه از دقت کافی برخـوردار نبـوده و امکـان تعمـیم و به کارگیری نتـایج حاصل از نمونه به جامعه وجود نداشته باشد. اما باید توجه داشت که همواره در نمونه‌های با حجم بزرگ نیز تضمینی برای اتخـاذ تـصمیم منطبـق بـر واقعیـت وجـود نـدارد. بیش از حد نیاز بودن اندازه نمونه همواره نقطه قوت مطالعه به شمار نمی‌آید. زیرا باعث از دست دادن منابع مالی، صرف وقت بیش از اندازه و ... می‌شود.

برای یافتن روش تعیین حجم نمونه نیاز است که محقق در ابتدا به چند سؤال پاسخ دهد:

1. نوع پژوهش توصیفی است یا تحلیلی؟
2. متغیر اصلی مورد پژوهش کمی است یا کیفی؟
3. مقدار شاخص پراکندگی چقدر است؟
4. محقق چقدر خطا را می‌پذیرد؟
5. چه میزان اطمینان و در صورت لزوم چه توانی برای نیل به هدف کافی است؟

انواع خطا در آمار

خطای نوع اول: اگر فرضیه صفر درست به اشتباه رد شود. احتمال خطای نوع اول را با α نمایش می‌دهند.

خطای نوع دوم: اگر فرضیه صفر نادرست به اشتباه پذیرفته شود. احتمال خطای نوع دوم را با β نمایش می‌دهند.

توان آزمون: احتمال رد کردن فرضیه صفر نادرست که برابر با $1-\beta$ است.

	فرض صفر درست	فرض صفر غلط
رد فرض صفر	خطای نوع اول α	توان آزمون $1-\beta$
پذیرش فرض صفر	خطایی رخ نداده $1-\alpha$	خطای نوع دوم β

محقق باید در پیش نویس طرح به موارد ذیل اشاره کند:

- ذکر آن هدف اختصاصی که ملاک محاسبه حجم نمونه قرار گرفته است.
- فرمول یا جدول مناسب برای محاسبه حجم نمونه متناسب با هدف اختصاصی مذکور.
- بیان منطق یا منبع اعداد و ارقامی که برای محاسبه حجم نمونه در فرمول قرار می‌گیرند.
- درصورت نیاز لحاظ کردن ضرایبی برای اطمینان از رسیدن به عدد محاسبه شده (پیش بینی میزان عدم مراجعه و عدم پاسخ و ...)

1- اندازه نمونه در مطالعات توصیفی

در مطالعات توصیفی، معمولاً هدف برآورد میانگین یک صفت کمی یا نسبت یک صفت کیفی است. پژوهشگر، در این نوع مطالعه می‌خواهد برآورد خود را با حداکثر فاصله معینی از واقعیت (دقت) و با اعتماد معینی از وجود پارامتر در آن فاصله (حدود اطمینان) به دست آورد.

الف) اندازه نمونه لازم برای برآورد میانگین یک صفت کمی در جامعه نامحدود با فرمول 1 محاسبه می‌شود:

برای محاسبه اندازه نمونه جهت برآورد یک میانگین می‌بایست برآوردی از اندازه انحراف معیار میانگین اندازه مورد نظر که می‌تواند از مطالعات قبلی یا مطالعه راهنما بدست آمده وجود داشته باشد.

$$n = \frac{z^2_{1-\frac{\alpha}{2}}\ \delta^2}{d^2}$$ (فرمول 1)

مثال 1: اگر انحراف معیار قد نوزادان جامعه 3 و بیشترین خطای قابل قبول یک سانتیمتر باشد. اندازه نمونه لازم برای برآورد میانگین قد نوزادان این جامعه با سطح اطمینان 95 درصد چقدر است (35).

مثال2: یک متخصص تغذیه در نظر دارد مطالعه‌ای برای برآورد متوسط مصرف پروتئین روزانه افراد نوجوان انجام دهد. با استفاده از مطالعات قبلی انحراف معیار مصرف پروتئین در جامعه مورد مطالعه حدود 20 گرم است. اگر حداکثر خطای برآورد 5 گرم و فاصله اطمینان 95٪ در نظر گرفته شود. حجم نمونه مورد نیاز چقدر است (62).

ب) اندازه نمونه لازم برای برآورد نسبت یک صفت کیفی در جامعه نامحدود با استفاده از فرمول 2 محاسبه می‌شود:

برای محاسبه اندازه نمونه می‌بایست برآوردی از اندازه **p** که می‌تواند از مطالعات قبلی بدست آمده وجود داشته باشد (گاه نیز می‌توان با توجه به تجربیات افراد متخصص در آن زمینه و یا مطالعه راهنما این اندازه را برآورد نمود). در صورتی که هیچ یک از راه‌های ذکر شده در بالا عملی نیست، p=0/05 فرض شود تا حداکثر تعداد نمونه ممکن بدست آید.

$$n = \frac{Z_{1-\frac{\alpha}{2}}^{2} P(1-P)}{d^2}$$ (فرمول2)

- همیشه باید با توجه به شاخص پراکندگی (انحراف معیار) خطا را تعیین کرد. در مورد برآورد یک میانگین ملاک خاصی برای تعیین خطا وجود ندارد اما برای برآورد یک نسبت معمـولاً خطا را بین 0/1 تا 0/25 نسبت مذکور در نظر می‌گیرند و روشن است که هرچه نـسبت مذکور به 0/1 نزدیکتر شود، مقدار dکوچک‌تر می‌شود و لذا حجـم نمونه محاسبه شده بیشتر خواهد شد.

- توجه به این مطلب لازم است که در صورتی که p=0/05، باید مقدار d با دقت و احتیاط انتخاب شود؛ زیرا فرض p=0/05 به خاطر ناآگاهی از حدود آن است.

مثال 1: محققی در نظر دارد نسبت افراد مبتلا به سردرد میگرنی را در شهر شیراز برآورد کند. با توجه به اطلاعات بدست آمده در پژوهش‌های قبلی این نسبت 5 درصد گزارش شده است. اگر محقق درصد اطمینان را 95 و خطای 20 درصد نسبت مذکور را در نظر بگیرد حجم نمونه چقدر خواهد بود (1825 نفر).

- **مثال 2:** پژوهشگری می‌خواهد نسبت ابتلا به گواتر را در یک شهرستان برآورد کند. مطالعات قبلی مقدار گواتر را 40٪ نشان داده است. تعداد نمونه را در سطح 95٪ اطمینان و با دقت 4٪ محاسبه کنید (576 نفر).

ضریب تصحیح جامعه محدود

در مواردی که جامعه مورد مطالعه کوچک (محدود) باشد، پس از محاسبه n با استفاده از فرمول‌های قبل حجم نمونه را محاسبه می‌کنیم و سپس با استفاده از ضریب زیر حجم نمونه اصلاح می‌شود. N حجم جامعه و n_o حجم نمونه اولیه بدست آمده با فرمول‌های مذکور است.

$$n' = \frac{n}{1+\frac{n}{N}}$$ (فرمول 3)

2- اندازه نمونه در مطالعات تحلیلی

در مواردی که پژوهش تحلیلی است یعنی محقق بدنبال یافتن اختلاف بین گروه‌ها با آزمون ارتباط بین متغیرها است. علاوه بر مقوله اطمینان به نتایج، توان آزمون مورد استفاده در کشف یا نشان دادن حقیقت نیز مطرح می‌شود. مقدار این توان معمولاً 80 درصد در نظر گرفته می‌شود.

الف- برآورد اندازه نمونه برای تشخیص اختلاف میانگین در دو جامعه مستقل:

$$n = \frac{\left(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}\right)^2 (\delta_1^2 + \delta_2^2)}{(\mu_1 - \mu_2)^2}$$

(فرمول 4)

مثال 1: در یک پژوهش محقق با هدف مقایسه دو روش آموزش می‌خواهد متوسط نمره کسب شده افراد را پس از آموزش با هم مقایسه کند. در پژوهش‌های پیشین بدست آمده است که نمره فراگیران با روش آموزشی اول 18 با انحراف معیار 3 و در روش دوم 16 با انحراف معیار 2/5 است. با 95 درصد اطمینان و 80 درصد توان آزمون چه حجم نمونه‌ای در کل لازم است؟ (30 نفر در هر گروه).

مثال 2: محققی قصد دارد نتایج کوتریموکسازول و روش کلاسیک را در درمان توکسوپلاسموز چشمی با هم مقایسه کند. با توجه به مطالعات قبلی میانگین درصد کاهش ضایعه با درمان کلاسیک 50 و با روش کوتریموکسازول 55 و انحراف معیار هر دو گروه برابر 6/5 در نظر گرفته شد. با 95 درصد اطمینان و 90 درصد توان آزمون چه حجم نمونه-ای در هر گروه لازم است (36 نفر در هر گروه).

ب- برآورد اندازه نمونه برای تشخیص اختلاف نسبت در دو جامعه:

$$n = \frac{\left(Z_{1-\frac{\alpha}{2}}\sqrt{2\bar{P}(1-\bar{P})} + Z_{1-\beta}\sqrt{P_1(1-P_1)+P_2(1-P_2)}\right)^2}{(P_1 - P_2)^2} \qquad \bar{P} = \frac{P_1 + P_2}{2}$$

(فرمول 5)

در این فرمول:

$\bar{P}$: متوسط دو نسبت P_1: نسبت در گروه اول P_2: نسبت در گروه دوم

یک فرمول راحتتر هم هست!

$$n = \frac{\left(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}\right)^2 [P_1(1-P_1) + P_2(1-P_2)]}{(P_1 - P_2)^2}$$ (فرمول 6)

مثال 1: محققی می‌خواهد در پژوهشی بین بیماران مبتلا به رفلاکس معده - مری و افراد سالم مصرف دخانیات را به عنوان یکی از علل مقایسه کند. در پژوهش‌های سابق میزان مصرف دخانیات در این بیماران 35 درصد و در افراد سالم 20 درصد می‌باشد. با 95٪ اطمینان و 80٪ توان آزمون چه حجم نمونه‌ای لازم است (138 نفر در هر گروه).

مثال 2: پژوهشگری می‌خواهد نتایج آناتومیک ویترکتومی اولیه و باکل اسکلرا در درمان جداشدگی شبکیه در بیماران دارای سابقه جراحی آب مروارید با هم مقایسه کند. با توجه به مطالعات قبلی درصد جداشدگی شبکیه بعد از 6 ماه پیگیری در گروه باکل اسکلرا 30٪ و در گروه ویترکتومی 40٪ در نظر گرفته شد. با 95 درصد اطمینان و 80 درصد توان آزمون چه حجم نمونه‌ای در هر گروه لازم است (125 نفر در هر گروه).

ج- برآورد اندازه نمونه برای زمانی که ضریب همبستگی در یک مطالعه مدنظر باشد:

(فرمول7) $\omega = \frac{1}{2} Ln \frac{1+r}{1-r}$ ← حدود اعتماد ضریب همبستگی

$$n = \frac{\left(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}\right)^2}{(\omega)^2} + 3$$

مثال 1: محققی قصد دارد همبستگی بین حرکات دیافراگم و شدت بیماری را در مبتلایان به COPD بسنجد. در مطالعات پیشین بدست آمده است که همبستگی بین حرکات دیافراگم و FVC در بیماران COPD 0/302 می‌باشد. با 95 درصد اطمینان و 80 درصد توان آزمون چه حجم نمونه‌ای در کل لازم است (84 نفر).

مثال 2: پژوهشگری می‌خواهد همبستگی بین کیفیت زندگی کاری و رضایت شغلی را در پرستاران بررسی نماید. بر اساس مطالعات قبلی این همبستگی 0/23 بدست آمده است. با 95 درصد اطمینان و 90 درصد توان آزمون چه حجم نمونه‌ای در کل لازم است (194 نفر).

جدول مورگان برای محاسبه حجم نمونه

در مواردی که واریانس جامعه یا درصد مورد نیاز را در اختیار نداشته باشید، می‌توان از جدول مورگان برای برآورد حجم نمونه استفاده کرد. این جدول حداکثر تعداد نمونه را در برآورد میزان شیوع نشان می‌دهد.

حجم نمونه	جامعه آماری	حجم نمونه	جامعه آماری	حجم نمونه	جامعه آماری
291	1200	140	220	10	10
297	1300	144	230	14	15
302	1400	148	240	19	20
306	1500	152	25	24	25
310	1600	155	260	28	30
313	1700	159	270	32	35
317	1800	162	280	36	40
320	1900	165	290	40	45
322	2000	169	300	44	50
327	2200	175	320	48	55
331	2400	181	340	52	60
335	2600	186	360	56	65
338	2800	191	380	59	70
341	3000	196	400	63	75
346	3500	201	420	66	80
351	4000	205	440	70	85
354	4500	210	460	73	90
357	5000	214	480	76	95
361	6000	217	500	80	100
364	7000	226	550	86	110
367	8000	234	600	92	120
368	9000	242	650	97	130
370	10000	248	700	103	140
375	15000	254	750	108	150
377	20000	260	800	113	160
379	30000	265	850	118	170
380	40000	269	900	123	180
381	50000	274	950	127	190
382	75000	278	1000	132	200
384	100000	285	1100	136	210

آشنایی با برخی از نرم افزارهای محاسبه حجم نمونه

منابع

1- ملک‌افضلی ح، مجدزاده ر، فتوحی ا، توکلی س. روش‌شناسی پژوهش‌های کاربردی در علوم پزشکی. بشردوست ن: نمونه‌گیری و اندازه نمونه. چاپ اول. تهران؛ انتشارات علوم پزشکی تهران، 1384: 316-295

2- حقدوست ع. آیا می‌خواهید مفهوم حجم نمونه و توان آزمون‌های آماری را دقیق‌تر درک کنید؟ مجله اپیدمیولوژی ایران. 1387؛ 5(1): 63-57

3- فلاح‌زاده ه. مبانی و روش‌های آمار زیستی. چاپ دوم. تهران، انتشارات سبحان، 1392: صفحه 11-206