

روش نمونه‌گیری تصادفی

گسترش مطالعات تجربی در زمینه مسایل اقتصادی-اجتماعی، امروزه بیش از پیش به کاربرد روشهای آمارگیری نمونه‌ای، وابسته است، چراکه انجام طرحهای تجربی در جوامع بزرگ فقط از طریق نمونه‌گیری ممکن است. در پاسخ این سؤال که جوامع مبتنی بر نمونه تا چه حد می‌تواند بازگوکننده حقیقت مسایل اجتماعی باشد، باید گفت که در اینجا کیفیت نمونه‌گیری مطرح است، و از اینرو، شناخت اصول نظری روش نمونه‌گیری و تطابق آن با شرایط اجرایی اهمیت پیدا میکند. موضوع بحث «دکتر جلیل دهقانی» استادیار دانشگاه تهران در این مقاله اصول نظری روشهای نمونه‌گیری تصادفی است.

مقدمه

روشهای نمونه‌گیری دلخواه و برنامه‌ریزی شده غالباً نتایج قابل‌تعمیمی بدست نمی‌دهد. دسترسی به نتایج مورد اطمینان، یا به عبارت دیگر قابل‌تعمیم به جامعه اصلی، بر اساس اصول آماری، اساساً از طریق نمونه‌گیری تصادفی ممکن است. حداعتماد نتایج جامعه نمونه فقط در روش نمونه‌گیری تصادفی به‌طور کمی قابل‌سنجش است، و این موجب برتری واقعی نمونه‌گیری تصادفی به نمونه‌گیری‌های گویای دیگر است. در نمونه‌گیری تصادفی نیز روشهای مختلفی در انتخاب نمونه‌ها بکار گرفته می‌شود. مهمترین روشهای نمونه‌گیری تصادفی عبارتند از:

- نمونه‌گیری تصادفی ساده (Einfache Zufallsauswahl)
- نمونه‌گیری تصادفی منظم (Systematische Auswahlverfahren)

نامه پژوهشکده، سال دوم،
شماره دوم: تابستان ۲۵۳۷

- نمونه گیری مختلط (Gemischte Auswahlverfahren)
- نمونه گیری قشر بندی شده (Geschichtete Auswahlverfahren)
- نمونه گیری خوشه ای (Flaechen-u. Klumpenstichprobe)
- نمونه گیری چند مرحله ای (Mehrstufige Auswahlverfahren)
- نمونه گیری مسیری (Random-(Route) Walk Sampling)
- نمونه گیری انباشته ای (Kumulationsverfahren)

نمونه گیری تصادفی ساده

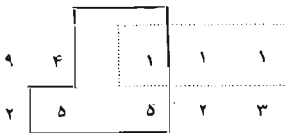
در این روش که نمونه گیری تصادفی نامحدود نیز نامیده می شود، کلیه عناصر جامعه اصلی برای انتخاب، شانس کاملاً یکسانی دارند.

ساده ترین راه برای انتخاب عناصر نمونه در این روش (۳) استفاده از الگوی صندوقی یا گردونه ای (Urnenmodell) است. در اینجا کلیه عناصر اصلی را در سرت امکان در یک صندوق یا گردونه می ریزند و پس از مخلوط کردن کامل این عناصر، تعداد نمونه های مورد نیاز را از آن به صورت قرعه بیرون می کشند. مادامی که نتوانیم سرت مورد تحقیق را در صندوق بریزیم، بکار گرفتن این روش عملاً مشکل خواهد بود. در چنین مواردی باید ابتدا عناصر جامعه اصلی را شماره گذاری کرده (مثلاً کلیه کتابفروشی های تهران را) و سپس شماره ها را در صندوق ریخته و نمونه گیری کرد. بدیهی است اگر تعداد عناصر جامعه اصلی زیاد باشد (مثلاً کلیه خانواده های دهقانی ایران) الگوی فوقی را نمیتوان بکار گرفت و باید از جداول خاصی بنام جداول اعداد تصادفی استفاده کرد. این جداول قبلاً از طریق مخلوط کردن دقیق اعداد و انتخاب کاملاً تصادفی آن تهیه شده است (۱). جهت شناخت روش استفاده از جداول اعداد تصادفی، به ذکر مثال می پردازیم:

از یک جامعه اصلی $N = 200$ باید به کمک جدول اعداد تصادفی ۱۰۰ نفر نمونه گیری کنیم. جامعه اصلی را کارمندان اداری یک دانشکده تصور می کنیم. این افراد را به دلخواه و به صورت اعداد سه رقمی از ۰۰۱ تا ۲۰۰ شماره گذاری کنیم. از طریق جدول اعداد تصادفی ۲۰ نمونه ($n = 20$) را بین اعداد ۱ تا ۲۰۰ انتخاب می کنیم. به عنوان مثال چند نمونه را میتوان از یک خانه (Block) ۵ نفره جدول انتخاب کرد (۱-صفحه ۳۸).

۷ ۰ ۸ ۳ ۰
 ۲ ۲ ۹ ۴ ۸
 ۴ ۱ ۶ ۴ ۲
 ۹ ۴ ۱ ۱ ۱
 ۲ ۵ ۵ ۲ ۳

بنست آوردن اعداد سه رقمی زیر ۲۰۰ از طریق خانه‌های ۵ رقمی جدول اعداد به دلخواه صورت می‌گیرد. از خانه فوق می‌توان اعداد ۱۵۵ و ۱۱۱ را به ترتیبی که



در بالا مشخص شده انتخاب کرد. به همین ترتیب و بسادگی می‌توان ۲۰ عدد سه رقمی را در یک جدول انتخاب کرد. و سپس بر اساس شماره‌ها عناصر نمونه را شناخت. بکار گرفتن این روش بخصوص وقتی مناسب است که عناصر جامعه اصلی خواه ناخواه قبلاً شماره‌گذاری شده باشند.

در روش تصادفی ساده، و بخصوص در الگوی صندوقی، انتخاب نمونه‌ها می‌تواند با برگشت به صندوق و یا بدون برگشت به صندوق صورت گیرد. انتخاب یکی از این دو روش بستگی به هدف تحقیق دارد که آیا یک عنصر جامعه می‌تواند دو یا چند بار مورد مطالعه قرار گیرد؟

فن نمونه‌گیری تصادفی. در روش نمونه‌گیری تصادفی ممکن است اشتباهات منظم خاصی رخ دهد، که میزان و جهت آن قابل تخمین نیست. به عنوان مثال در مخلوط کردن ورق بازی ثابت شده است که هیچ یک از نحوه‌های متداول بر زدن، موجب نا-مرتبی کامل کارتها نمی‌شود (۲). این آگاهی در مورد کلیه روش‌های نمونه‌گیری تصادفی حائز کمال اهمیت است. اینکه آیا خصوصیات فیزیکی جامعه اصلی، نمونه‌گیری تصادفی ساده را ممکن می‌کند یا نه از مسایل قابل توجه است. به عنوان مثال نمونه‌گیری از جامعه آماری شراب، اسید، کره، چای، معدن و غیره غالباً با تبادل نظر متخصصان و آمارگران، ممکن است.

نون نمونه‌گیری باید از اینگونه اشتباهات غیرقابل کنترل جلوگیری کند.

این کار در دو جهت صورت می‌گیرد: از یک طرف با منظم کردن عناصر جامعه اصلی و از طرف دیگر با مشخص کردن صفاتی که انتخاب نمونه‌ها با اتکا به آنها صورت می‌گیرد. شرط اصلی، تنظیم لیست، پرونده و یا هرگونه مشخصات دیگر واحدهای آماری است، به طوری که عناصر جامعه اصلی قابل شناخت و صفات مورد پژوهش قابل تشخیص باشد. در هر یک از روشهای نمونه‌گیری تصادفی باید مدارک اصلی بدون استثنا بطور منظم با وضع فعلی جامعه اصلی، مطابقت کند. در غیر اینصورت انتخاب تصادفی کاملاً دقیق، ناممکن است. تنظیم عناصر مورد تحقیق از طریق جدول تصادفی اعداد و در نمونه‌گیری منظم دارای اهمیت فوق‌العاده زیادی است. تنظیم لیست یا پرونده‌ها، باید به گونه‌ای باشد که بخت انتخاب واحدهای مورد پژوهش یکسان شود. به عنوان مثال اگر هدف ما انتخاب چند خانوار در یک بنگاه تولیدی باشد و در این بنگاه تولیدی چند نفر از یک خانواده مثلاً پدر، مادر و پسر فعالیت داشته باشند، یا به عبارت دیگر چندین واحد انتخابی را شامل بشوند، در اینصورت بخت کلیه خانوارها در انتخاب یکسان نیست.

نمونه‌گیری تصادفی منظم

برای انتخاب یک جامعه نمونه ۴۰۰ نفری از یک جامعه اصلی چند میلیون نفری، باید اولاً کلیه آدرسها و اسامی، شماره‌گذاری شود و ثانیاً از میان جامعه آماری ۴۰۰ مورد مثلاً از طریق جدول اعداد تصادفی، انتخاب شود. بدیهی است که روش نمونه‌گیری تصادفی ساده در چنین مواردی با توجه به میزان کار، زمان مورد احتیاج و دقت، روشی نامناسب است. از این رو روش نمونه‌گیری تصادفی منظم مورد توجه خاص محققان قرار گرفت. انتخاب نمونه‌ها در این روش به طرق مختلف صورت می‌گیرد که عبارتند از:

۱- نمونه‌گیری براساس نسبت بین تعداد جامعه اصلی و جامعه نمونه

۲- نمونه‌گیری براساس رقم آخر شماره ردیف عناصر جامعه اصلی

۳- نمونه‌گیری براساس حرف اول اسامی عناصر جامعه اصلی و تاریخ تولد آنها

نمونه‌گیری براساس نسبت بین جامعه اصلی و جامعه نمونه. در این روش نسبت

بین جامعه اصلی و جامعه نمونه را حساب می‌کنیم $(\frac{N}{n} - k)$ ، و سپس از طریق

یک جدول اعداد تصادفی یا ابزار انتخاب تصادفی دیگر، یک عدد i را که بین ۱ و K

قرار دارد به طور تصادفی انتخاب می‌کنیم. کلیه واحدهای جامعه اصلی در ردیف

$i, i+k, i+2k, i+3k, \dots$ جامعه نمونه ما را تشکیل می‌دهد.

در مواردی که نام و آدرس جامعه اصلی قبلاً روی نوارهای مغناطیسی کامپیوتر (به منظور استفاده های دیگر) ضبط شده باشد، بکار گرفتن این روش بسیار ساده خواهد بود. در اینجا کافی است که K و شماره اولین عنصر را مشخص کنیم و سپس از طریق يك برنامه کامپیوتری نام و آدرس عناصر دیگر را، که در فاصله $i+k$ ، $i+2k$ و غیره قرار دارند بدست آوریم. ثابت بودن فاصله بین دو عنصر انتخابی، ممکن است در بعضی موارد موجب اشتباه زیادی بشود. به عنوان مثال ممکن است در پژوهش ساخت آگهی-های روزانه ای که هر روز در هفته انتشار می یابد اشتباه زیر اتفاق بیفتد. اگر نمونه ها در اینجا به ترتیب گردش انتشار و یا به عبارت دیگر روزنامه هر ششمین روز انتخاب شود، در این صورت ظاهر فقط روزنامه يكروز معینی از هفته، مثلاً روزنامه پنجشنبه به عنوان نمونه انتخاب می شود. بدیهی است که ساخت آگهی های روزنامه پنجشنبه قابل تعمیم به تمام هفته نیست (گویا نیست).

در پرونده های اداری نیز ممکن است دوره هایی بر اساس يك صفت معین وجود داشته باشد. به عنوان مثال در پرونده متاهلین، نام مرد غالباً اول و نام زن دوم نوشته می شود. در اینگونه موارد کلیه اعداد فرد را مردها و کلیه اعداد زوج را زنهای تشکیل می دهند. اگر K يك عدد زوج باشد، نمونه های ما با فقط مرد هستند یا فقط زن، و اگر K فرد باشد زن و مرد به نسبت مساوی نمونه ها را تشکیل خواهند داد.

بنابراین باید توجه کامل داشت که آیا دوره های خاصی که برای صفات مورد پژوهش معنی دار است، وجود دارد یا نه. در صورت وجود طول هر دوره نباید با طول فاصله دو نمونه متوالی یکسان باشد. به اشتباهاتی که در فوق به عنوان مثال ذکر شد می توان با کمی دقت و شناخت به اصول اساسی نمونه گیری به سادگی پی برد، لکن در تحقیقات اجتماعی مواردی وجود دارد که خطر اشتباه به سادگی قابل شناخت نیست.

نمونه گیری بر اساس رقم آخر شماره ردیف عناصر جامعه اصلی. در این روش عناصری به عنوان نمونه انتخاب می شوند که رقم کارت آنها مثلاً به عدد ۲۰۹ ختم بشود. هر يك از اعداد ۹۰۲ جداگانه ۱۰ درصد از عناصر جامعه اصلی را تشکیل می دهند. جهت انتخاب ۴ درصد از جامعه اصلی، کافی است که کلیه واحدهای آماری که رقم آخر شماره آنها مثلاً به ۷۳۰۴۸، ۲۱، ۲۱، ۲۱ ختم می شود، انتخاب کنیم. انتخاب نمونه ها از طریق این روش مستلزم داشتن ردیف کامل اعداد ۱ تا N نیست، بلکه در اینجا کافی است که رقم آخر شماره کارتها به طور نسبتاً مساوی بین اعداد تقسیم شده باشد (۳). اعداد انتخابی مانند ۷۳، ۴۸، ۲۱، ۲۱، ۲۱ به طور تصادفی انتخاب می شوند.

نمونه‌گیری براساس حرف اول اسامی عناصر جامعه اصلی و تاریخ تولد آنها

نمونه‌گیری براساس حروف و تاریخ تولد روشی است که شبیه به روش نمونه‌گیری تصادفی ساده می‌باشد. این روش بیشتر در مواقعی که افراد واحدهای مورد مطالعه هستند، بکار گرفته می‌شود. در اینجا کلیه افرادی انتخاب می‌شوند که نام خانوادگی آنها با یک حرف کاملاً مشخصی شروع شده و در تاریخ معینی، متولد شده‌اند. بدیهی است که می‌توان چندین حرف و یا تاریخ تولدهای مختلفی را جهت انتخاب نمونه‌ها در نظر گرفت که به تعداد نمونه‌ها بستگی داشته باشد. بکار گرفتن این روش بر اساس این فرضیه است که بین حرف اول نام خانوادگی و تاریخ تولد از یکطرف، و پدیدارهای آماری مورد نظر از طرف دیگر، همبستگی وجود ندارد. به عنوان مثال تاریخ تولد تأثیری روی میزان درآمد ندارد، یا بین قد مردهای ۳۰ تا ۳۵ سال که نام آنها با «ب» شروع می‌شود با قد مردهای ۳۰ تا ۴۵ سال که نام آنها با «د» شروع می‌شود رابطه‌ای وجود ندارد:

بکار گرفتن روش مزبور، درعین حال مستلزم آن است که بدانیم تقریباً چقدر صد از نام‌های خانوادگی با هر یک از حروف شروع می‌شود. براساس این شناخت می‌توان تخمین زد که چند درصد، و یا مطلقاً چند نفر از مثلاً کل کارکنان یک بنگاه تولیدی باید انتخاب بشوند. تعداد کارکنان و به‌طور کلی تعداد عناصر جامعه اصلی را می‌توان از طریق راههای دیگر تهیه کرد.

این روش دارای دو جنبه مثبت است: اولاً فقط یک قسمت از عناصر مورد آزمایش قرار می‌گیرند و ثانیاً زمان استخراج پرسشها به‌طور قابل ملاحظه‌ای کاهش می‌یابد. روش نمونه‌گیری بر اساس حروف باید به مسایل خاصی توجه داشت. به عنوان مثال حروف اول اسامی دهقانان مهاجر یک منطقه غالباً با اسامی افرادی که متولد شهر هستند متفاوت است و انتخاب یک حرف ممکن است باعث تحریف نتایج پژوهشهایی که در صورت می‌گیرد، بشود. از اینرو بهتر است که چند حرف را برای انتخاب نمونه‌ها در نظر بگیریم.

نمونه‌گیری منظم همانطور که مشاهده می‌شود، عملاً و به‌سادگی قابل استفاده است، ولی این سؤال نیز مطرح است که آیا این روش می‌تواند جایگزین روش نمونه‌گیری تصادفی ساده باشد یا نه؟ روش نمونه‌گیری منظم هنگامی تصادفی است که پرونده‌ها به‌طور منظم رده‌بندی نشده باشد، که عملاً کمتر با آن مواجه می‌شویم. لیست اسامی و آدرس و پرونده‌ها معمولاً براساس یک صفت مشخص تنظیم می‌شود. مثلاً به ترتیب حروف الفبا و یا براساس مثلاً شماره عضویت و غیره. بدیهی است که نمی‌توان به منظور نمونه‌گیری تصادفی منظم، نظم پرونده‌ها را برهم زد. انتخاب شماره اول به‌طور تصادفی هرچند بسیار مهم است، لکن متضمن انتخاب تصادفی بقیه نمونه‌ها در روش منظم نیست.

گفتیم که از دید نظری کلیه نمونه‌های ممکن به بزرگی n باید بخت متساوی در انتخاب داشته باشند. در نمونه‌گیری منظم فقط تعداد K از جامعه نمونه به بزرگی n از بخت متساوی برخوردارند .

در اینگونه نمونه‌گیری انتخاب آدرسهای بعدی همیشه تابع آدرس قبلی است، در صورتیکه در یک نمونه‌گیری کاملاً تصادفی هر یک از آدرسها باید به‌طور آزاد انتخاب بشود. نمونه‌گیری مزبور این جنبه مثبت را نسبت به روش نمونه‌گیری ساده دارد که تمام جامعه اصلی را زیر پوشش قرار می‌دهد. از این طریق درجه دقت نتایج افزایش می‌یابد. در نمونه‌گیری منظم هر چند انتخاب عناصر کاملاً تصادفی نیست ، لکن این روش عملاً زیاد و با نتایج رضایت بخش بکار گرفته می‌شود .

نمونه‌گیری مختلط

نمونه‌گیری مختلط ترکیبی است از نمونه‌گیری تصادفی و نمونه‌گیری برنامه‌ریزی شده که در پژوهشهای تجربی اقتصادی و اجتماعی غالباً بکار گرفته می‌شود. در زیر مهمترین انواع این روش را مورد بحث قرار می‌دهیم :

نمونه‌گیری قشر بندی شده . هدف نمونه‌گیری در پژوهشهای اقتصادی- اجتماعی ، دسترسی به نتایج نسبتاً دقیق است . از اینرو باید از طریق انتخاب روش نمونه‌گیری درجه اشتباه را بدون افزایش هزینه تحقیق به حداقل ممکن رساند . یا به عبارت دیگر از طریق بکار گرفتن یک روش نمونه‌گیری ساده‌تر با تعداد نمونه‌های مساوی باید نتایج دقیقتری بدست آوریم. در پژوهشهای اقتصادی - اجتماعی ممکن است که جامعه اصلی مورد نظر ، از نظر صفات معینی بسیار ناهمگن باشد. اگر پراکندگی جامعه اصلی (σ^2) زیاد باشد ، پراکندگی جامعه نمونه (σ^2) نیز زیاد خواهد بود. هر چه پراکندگی جامعه آماری زیادتر باشد ، حد اعتماد $(\bar{x} \pm t \frac{S}{\sqrt{n}})$ زیادتر (۳) و در نتیجه

دقت نتایج کمتر خواهد بود ، و نمونه‌های مختلف نتایج متفاوتی خواهد داشت . روش نمونه‌گیری قشر بندی شده بهترین روشی است که در اینگونه موارد بکار گرفته می‌شود. در روش قشر بندی ، جامعه اصلی بر اساس صفات مشخصی گروه بندی می‌شود. هر یک از گروهها را یک قشر (Strata) می‌نامند (۴) . قشرها نسبت به صفت یا صفاتی که بر اساس آن قشر بندی شده‌اند در درون خود همگن هستند .

پدیدهی است هر چه جامعه اصلی همگن‌تر باشد ، یا به عبارت دیگر واریانس کمتری داشته باشد ، تعداد نمونه‌های لازم برای رسیدن به یک درجه اطمینان (t) معین ،

کمتر خواهد بود. چون با تعداد نمونه های معینی از طریق قشر بندی درجه اطمینان در مقایسه با انتخاب تصادفی ساده افزایش می یابد، بنابراین روش قشر بندی شده نسبت به روشهای دیگر در موارد خاص مؤثرتر خواهد بود:

از طریق قشر بندی جامعه اصلی پراکندگی درونی کاهش می یابد و برعکس پراکندگی خارجی به صورت جدید ظاهر می شود، که در آن جوامع کوچک تری، براساس صفات مناسب تشابه کمتری دارند. هرچه پراکندگی خارجی میان قشرها زیادتر و پراکندگی درونی آنها کمتر باشد، روش قشر بندی شده مؤثرتر خواهد بود. همچنین هرچه تعداد قشرها زیادتر باشد درجه اشتباه نیز کمتر خواهد بود. تعداد نمونه هایی که در این روش انتخاب می شود بیشتر متأثر از واریانس درونی است (۵). از اینرو قشر بندی مطلوب جامعه اصلی، مستلزم شناخت دقیق واریانس قشرهاست که براساس یک تحقیق مقدماتی و یا آمارهای موجود، قابل محاسبه است. محاسبه واریانس کل جامعه اصلی برای مشخص کردن تعداد نمونه ها در یک جامعه قشر بندی شده به ترتیبی خاص محاسبه می شود، که در زیر به آن اشاره خواهیم کرد.

اگر در یک پرسشگری عمومی، جامعه اصلی را مثلاً براساس صفت جنس قشر بندی کنیم، و فرض کنیم که جامعه اصلی شامل ۴۷ درصد مرد و ۵۳ درصد زن می باشد، در این صورت تعداد مرد و زن را در جامعه نمونه مشخص کرده ایم. از این طریق نقش تصادفی صفت جنس محدود شده و یا به اصطلاح آماری زیر کنترل است. براساس این محدودیت درجه اشتباه که فقط اثر تصادفی بودن را روی ترکیب جامعه نمونه می سنجد، کاهش می یابد (۵).

روش قشر بندی براساس اندیشه های فوق بکار گرفته می شود. در این روش همانطور که قبلاً گفتیم جامعه اصلی به قشرهای همگن مختلفی، تقسیم می شود و سپس از هر قشر تعدادی نمونه انتخاب می شود. نتایجی که براساس تمام نمونه ها بدست می آید، برای تعیین پارامتر مورد مطالعه جامعه اصلی بکار گرفته می شود. در روش نمونه گیری قشر بندی شده دو سؤال زیر مطرح است:

– جامعه اصلی را براساس کدام صفت یا صفات باید قشر بندی کرد؟

– کل جامعه نمونه n را به چه نسبتی باید بین قشرها تقسیم کرد؟

صفات قشر بندی. بکار گرفتن نمونه گیری قشر بندی شده مشکل تر از نمونه گیری تصادفی ساده است. مشکلات قشر بندی با تعیین صفات آغاز می شود. پراکندگی این صفات را باید در جامعه اصلی شناخت. نوع و میزان همبستگی این صفات را با صفات مورد پژوهش باید دانست. بین صفات قشر بندی و صفات مورد پژوهش باید همبستگی زیادی وجود داشته باشد.

در قشر بندی عملاً تعدادی کم ، مثلاً يك يا دو صفت را اساس کار ، قرار می دهند ، لکن هر چه قشرها بر اساس چندین صفت همگن تر باشند اثر قشر بندی زیادتر خواهد بود . صفاًتی را که غالباً بر اساس آن قشر بندی می کنند عبارتند از :

مکان جغرافیایی - محل سکونت (شهری و روستایی) - قدرت خرید (حقوق ، درآمد) - وضع اجتماعی - بزرگی واحد آماری (تعداد خانوار وغیره) - مذهب - سواد - سن - جنس - شغل . و یا در زمینه واحدهای اقتصادی : نوع سازمان اقتصادی - نوع تولید (تولید و فروش) - بزرگی کارگاه یا بنگاه (سطح زیرکشت ، تعداد کارکنان ، تعداد دام) - میزان فروش .

در قشر بندی باید به نکات و مسایل زیر توجه کامل داشت :

اول - تعداد عناصر هر قشر (N_i) باید مشخص شود .

دوم - مجموع عناصر قشرها باید با جامعه اصلی N برابر باشد ، یا به عبارت دیگر $N = N_1 + N_2 + \dots + N_m$.

سوم - قشرها باید کاملاً از همه متمایز باشند . به عبارت دیگر هر واحد آماری باید متعلق به يك قشر باشد .

چهارم - جامعه اصلی را می توان بر اساس يك يا چند صفت قشر بندی کرد ، مثلاً بر اساس جنس و درآمد .

پنجم - بعضی مواقع قشرها به طور خود کار تشکیل داده می شود ، مثلاً هنگامی که شمارش افراد در مکان های مختلف و در زمانهای متفاوت صورت بگیرد . چنین وضعی هنگامی پیش می آید که نمایندگان فروش يك کالا در محل های مختلف ، آمار گیری معینی را انجام دهند . در استخراج و ارزشیابی این گونه تحقیقات باید قوانین نمونه گیری قشر بندی شده را رعایت کرد تا بتوان به نتایج صحیح تری دست یافت .

بعد از قشر بندی جامعه اصلی ، اطلاعات زیر حائز اهمیت هستند :

۱ - بزرگی مطلق و نسبی قشرها .

۲ - پراکنندگی نسبی و مطلق واحدهای آماری هر قشر ،

۳ - هزینه نسبی آمار گیری و پرمشگری هر قشر .

تعداد نمونه های هر قشر . بعد از تعیین صفات قشر بندی و مشخص کردن تعداد

قشرها (m) باید اندیشید که کل جامعه نمونه (π) را چگونه باید بین قشرها (π_i)

تقسیم کرد . هدف در تقسیم جامعه بین قشرها دسترسی به ارزش میانگین جامعه اصلی

(\bar{X}) با حدا اعتماد محدودی است (۲) . تقسیم جامعه نمونه نقش بسیار مهمی در نمونه

گیری قشر بندی شده برعهده دارد ، زیرا واریانس s^2_x را ، که بر اساس آن حد اعتماد

تعیین میشود، تحت تأثیر قرار می‌دهد.

جامعه نمونه را میتوان اصولاً نسبت به بزرگی قشرها (N_1, N_2, \dots, N_m) به صورت متناسب (Proportional) یا نامتناسب (Disproportional) تقسیم کرد.

تقسیم متناسب بر اساس فرمول $\frac{n_i}{n} = \frac{N_i}{N}$ صورت می‌گیرد. در این فرمول

$i=1/2, \dots, m$ تعداد گروههای جامعه نمونه است. به عبارت دیگر نسبت بین نمونه‌های هر قشر، برابر با نسبت بین قشرهای جامعه اصلی است. تقسیم متناسب جامعه نمونه بین قشرها موجب گویایی کامل جامعه نمونه می‌باشد. بخت انتخاب کلیه عناصر جامعه اصلی در تقسیم متناسب یکسان است. نمونه‌گیری قشر بندی شده با تقسیم متناسب جامعه نمونه بین قشرها، معادل نمونه‌گیری تصادفی ساده می‌باشد.

تقسیم نامتناسب زمانی بکار گرفته می‌شود که قشرها با هم برابر نباشند و این ناهمبندی در نمونه‌گیری نیز مورد توجه قرار گیرد. هنگامی که از قشرهای کوچکتر تعداد نسبتاً زیادی نمونه لازم باشد، باید تقسیم نامتناسب را بکار گرفت.

تقسیم نامتناسب به صورتهای مختلف انجام می‌گیرد. ساده‌ترین روش، تقسیم جامعه نمونه بین هر یک از قشرها به مقدار ثابت است. اگر تعداد قشرها K باشد و تعداد نمونه‌ها n ، در این صورت تعداد نمونه‌های هر قشر برابر با $\frac{n}{K}$ خواهد بود یا به عبارت دیگر

$$n = \sum_{i=1}^K n_i = K n_i$$

نمونه‌گیری کنترل شده (۲) نیز نامیده می‌شود، غالباً هنگامی بکار می‌رود، که هدف پژوهش مقایسه ارزش میانگین با ارزشهای هر یک از قشرها باشد. در چنین مواردی نیاز به تعمیم نتایج به کل جامعه آماری نیست. نتایج نمونه‌گیری کنترل شده اصولاً هنگامی رضایت بخش است که واریانس کلیه قشرها (σ_i^2) با هم برابر باشد. در تقسیم نامتناسب « اصل انتخاب کاملاً تصادفی » دقیقاً رعایت نمی‌شود.

در پژوهشهای تجربی جامعه نمونه را به صورت مطلوب نیز بین قشرها تقسیم می‌کنند. در تقسیم متناسب برای محاسبه S^2 (و از این طریق تخمین دقت) فقط تفاوت واریانس‌های هر قشر (σ_i^2) مورد توجه قرار می‌گیرد، لکن در نمونه‌گیری مطلوب تعداد عناصر نمونه هر قشر n_i نیز در نظر گرفته می‌شود. تعداد نمونه‌های هر قشر n_i ($i=1/2, \dots, m$) بر این شرط تقسیم می‌شود، که پراکندگی S^2 تقسیم نمونه‌ها به حداقل برسد. تقسیم مطلوب غالباً بر اساس هزینه تحقیق صورت می‌گیرد. در این

حالت باید نمونه‌ها با توجه به مسایل فوق طوری انتخاب شوند که بتوان با هزینه‌معین، تحقیق مورد نظر را انجام داد. از طریق نمونه‌گیری مطلوب می‌توان در پژوهش‌هایی که اصولاً پرهزینه هستند، به قشرهای خاصی بیشتر توجه کرد. از این طریق باید هزینه تحقیق کاهش یابد.

اگر از i قشر z واحد مورد مطالعه قرار گیرد در این صورت هزینه C_i ایجاد می‌شود. هزینه کل در چنین حالتی به صورت زیر خواهد بود (۶):

$$K = \sum_{i=1}^m n_i C_i$$

تعداد نمونه‌های هر قشر (n_i) از طریق تساوی زیر بدست می‌آید:

$$V\{X\} = \sum_{i=1}^m \left(\frac{N_i}{N} \right)^2 \frac{1}{n_i} \left(\frac{N_i}{n_i} - 1 \right) \sigma_i^2 = \text{Min}$$

واریانس جامعه اصلی S بر اساس پارامترهای C_i و σ_i و N_i به صورت زیر بدست می‌آید:

$$S = \sum_{i=1}^m \frac{N_i}{N} \sigma_i / C_i$$

بنا بر این هنگام نمونه‌گیری باید قشری را (i) زیادتر در نظر گرفت، که نسبت آن بزرگتر، و پراکندگی درونی آن (σ_i) نیز بیشتر باشد. قشرهایی که در آن هزینه مطالعه (C_i) هر واحد زیادتر باشد کمتر مورد توجه قرار می‌گیرند. در تقسیم بندی مطلوب کوچکترین واریانس $V\{X\}$ به صورت زیر بدست می‌آید:

$$V\{X\} = V_{\text{OPT}} = \frac{i}{N^2} \sum_{l=1}^m N_l \left(\frac{N_l}{n_l} - 1 \right) \sigma_l^2$$

$$= \frac{i}{N} \left\{ \sum_{i=1}^m \frac{N_i^2}{N n_i} \sigma_i^2 - \sum_{i=1}^m \frac{N_i}{N} \sigma_i^2 \right\}$$

از طریق این فرمول تعداد نمونه‌های هر قشر قابل محاسبه است (۶).

تعداد نمونه‌ها را می‌توان از طریق روش «نی‌مان» (Neymann) نیز بین قشرها تقسیم کرد. تقسیم نمونه‌ها در این روش بر اساس حداقل واریانس صورت می‌گیرد. به عبارت دیگر نمونه‌ها به طریقی بین قشرها، تقسیم می‌شوند که واریانس جامعه نمونه نسبت به ارزش متوسط جامعه اصلی به حداقل برسد. تعیین تعداد نمونه‌ها از طریق روش «نی‌مان» و محاسبه واریانس جامعه قشر بندی شده بر اساس فرمولهای زیر صورت می‌گیرد (ع):

پراکندگی متوسط درونی (Within) قشرها که در آن پراکندگی هر قشر با $\frac{N_i}{N}$ توزین می‌شود به صورت زیر بدست می‌آید (ع).

$$\sigma_w = \sum_{i=1}^m \frac{N_i}{N} \sigma_i$$

در این صورت تعداد نمونه‌های هر قشر از رابطه زیر بدست می‌آید:

$$\frac{n_i}{N} = \frac{n}{N} \frac{\sigma_i}{\sigma_w}$$

واریانس که بر اساس تعداد نمونه‌های فوق بدست می‌آید، در صورتی که $N - 1 \approx N_i$:

باشد به صورت فرمول زیر خواهد بود:

$$S_{Neym} = \sum_{i=1}^m \frac{N_i}{N} \left(\frac{\sigma_w}{n} - \frac{\sigma_i}{N} \right) \sigma_i$$

$$S_{Neym} = \frac{\sigma_w}{n} \sum_{i=1}^m \frac{N_i}{N} \sigma_i - \frac{1}{N} \sum_{i=1}^m \frac{N_i}{N} \sigma_i^2$$

$$S_{Neym} = \frac{1}{n} \sigma_w^2 - \frac{1}{n} \sigma_m^2$$

نمونه‌گیری خوشه‌ای

در روش نمونه‌گیری فوق واحدهای انتخابی، واحدهای تحقیقی نیز هستند. گاهی اوقات بر اثر شرایط خاصی، مترون به صرفه‌است که از اصل برابری واحدهای انتخابی و تحقیقی چشم‌پوشی کنیم. یا به عبارت دیگر، قبل از نمونه‌گیری واحدهای تحقیقی، واحدهای آماری دیگری که شامل عناصر تحقیق هستند انتخاب کنیم. این روش مخصوص هنگامی مناسب است که اولاً واحدهای جامعه اصلی از نظر جغرافیایی پراکنده باشند، و ثانیاً صورت، پرونده یا هیچگونه نشانه دیگری از آنها، در اختیار نباشد.

تهیه صورت اسامی، آدرس و غیره برای جوامع بزرگ، مثلاً جامعه مصرف‌کننده تهران بسیار طولانی و گران است. چنین هزینه‌های گزاف و اتلاف وقت را می‌توان از طریق انتخاب گروهی از جامعه اصلی کاهش داد. اینگونه واحدهای آماری بزرگتر را خوشه (Klumpen) می‌نامند (۴). این گروه‌ها یا خود بخود وجود دارد (مانند کالای بسته-بندی شده در جعبه‌های کوچکتر) و یا اینکه می‌توان آنها را به سادگی تشکیل داد (مثلاً هر ۷۰ نفر از یک صورت). گروههای شبیه گروه اول را گروههای طبیعی و گروههای نوع دوم را مصنوعی می‌نامند. در روش نمونه‌گیری خوشه‌ای باید محل واحدهای تحقیق در هر گروه کاملاً مشخص باشد. نمونه‌گیری گروهی در دو مرحله صورت می‌گیرد:

ابتدا چند گروه از بین کلیه گروههای جامعه اصلی به طور تصادفی انتخاب می‌شوند و سپس واحدهای نمونه را در گروههای نمونه انتخاب می‌کنیم. به عنوان مثال اگر بخواهیم درباره وضع اقتصادی - اجتماعی اجاره نشینهای تهران پژوهشی بر اساس نمونه‌گیری گویا انجام دهیم، عملاً به این مشکل برخورد می‌کنیم که هیچگونه آمار و نشانه‌ای از این گروه در اختیار نداریم، تا بتوانیم براساس آن جامعه نمونه را انتخاب کنیم. برای این منظور جامعه نمونه را به طریق زیر انتخاب می‌کنیم: شهر را روی نقشه به ناحیه‌های بزرگی تقسیم می‌کنیم، ناحیه‌ها را شماره‌گذاری می‌کنیم، از بین این ناحیه‌ها ده درصد را به طور تصادفی انتخاب می‌کنیم، و در هر یک از این ناحیه‌ها، آمار-گیری دقیقی در باره اجاره‌تشینها انجام می‌دهیم.

روش مزبور را می‌توان حتی هنگامی که نشانه ویست اسامی اجاره نشینان نیز در اختیار است، به دو علت زیر بکار گرفت:

الف - چون پرسشگری و پژوهش در محدوده جغرافیایی معینی صورت می‌گیرد، بنابراین از یکطرف از هزینه‌های ایاب و نهاب و از طرف دیگر از اتلاف وقت کاسته می‌شود.

ب - آمارها و نشانه‌ها غالباً کامل نیست و تغییر در جامعه اصلی پویایی دائمی است.

در بکار گرفتن روش نمونه‌گیری خوشه‌ای باید آگاه بود که گردهم‌آیی واحدهای تحقیقی همگن در یک یا چند گروه (ناحیه) موجب عدم شناخت جامعه اصلی براساس نتایج تحقیق خواهد بود. به عنوان مثال اگر برای یک پرسشگری گویا در سازمان مرکزی یک بنگاه تولیدی، که در آن محل کار هر یک از کارکنان متناسب با سطح مدیریت آنها بین طبقات مختلف یک ساختمان مشخص شده باشد، و هر طبقه به عنوان گروه انبوهی انتخاب بشود در این صورت به طور یقین نتایجی که از جامعه نمونه بدست خواهد آمد، تصویری

تعریف شده از سطح تحصیلات، درآمد، طرز تفکر و غیره، درباره جامعه اصلی بدست خواهد داد.

نمونه گیری قشربندی شده یا نمونه گیری خوشه‌ای يك تفاوت اساسی دارد: در نمونه گیری قشربندی شده از کلیه قشرها نمونه گیری به عمل می‌آید، در صورتیکه در نمونه گیری خوشه‌ای فقط تعدادی از گروهها برای تحقیق انتخاب می‌شوند و برعکس نمونه گیری قشربندی شده، در اینجا کلیه عناصر تحقیق موجود در گروه، مورد مطالعه قرار می‌گیرند (۳). از اینرو واریانس نمونه گیری خوشه‌ای معمولاً نا محدودتر از پراکندگی در نمونه گیری تصادفی است، و در نمونه گیری قشربندی شده، حد اقل در مورد صفت قشربندی حالتی عکس این وجود دارد.

نمونه‌گیری چند مرحله‌ای

اگر در گروههایی که از طریق روش نمونه گیری خوشه‌ای انتخاب شده‌اند مجدداً نمونه گیری به عمل آید، روش نمونه گیری چند مرحله‌ای خواهد بود. در پژوهشهای اقتصادی - اجتماعی انتخاب عناصر نمونه باید شالباً در چند نوبت صورت گیرد. به عنوان مثال اگر واحد تحقیق خانوار دهقانی باشد، باید به ترتیب زیر عمل کرد:

- در مرحله اول باید از بین مثلاً کلیه دهستانها چند دهستان را در سطح کشور انتخاب کرد،

- در مرحله دوم باید چند ده را در دهستانهای نمونه انتخاب کرد،

- در مرحله سوم باید در هر ده، تعدادی خانوار انتخاب کرد.

بدیهی است که در هر مرحله، واحد انتخابی جدید کوچکتر از واحد انتخابی قبلی است.

انتخاب نمونه‌ها در هر مرحله با کلیه مراحل می‌تواند از طریق نمونه گیری تصادفی ساده، نمونه گیری قشری یا خوشه‌ای صورت گیرد. اگر روش نمونه گیری خوشه‌ای بکار گرفته شود، محاسبه درجه اشتباه نمونه گیری دشوار می‌شود.

از طریق روش نمونه گیری چند مرحله‌ای می‌توان با هزینه کم و عتلائی ترکار کرد. به این موضوع نیز باید توجه داشت که واحدهای انتخابی آخرین مرحله لزوماً نباید با واحدهای تحقیق برابر باشد. آخرین واحد انتخابی می‌تواند خوشه‌ای و شامل واحدهای تحقیق باشد.

نمونه‌گیری مسیری

روش نمونه گیری مسیری نوعی نمونه گیری خوشه‌ای است که در آن ابتدا

حیله‌هایی از تمام شهر انتخاب می‌شوند، و سپس در آن مسیر پرسشگر به ترتیبی که انتخاب خانه‌ها به طور تصادفی صورت گیرد مشخص می‌شود. به عنوان مثال به پرسشگر گفته می‌شود که در چهارراه X اولین خیابان سمت شمال را انتخاب کن، در این خیابان در نده رو دست راست حرکت کن، وارد سومین خیابان دست راست شو، در این خیابان بچمین خانه دست راست را انتخاب کن... و غیره. در این روش کوچه‌های بن‌بست، رلک‌ها و غیره نیز در نظر گرفته می‌شوند (۷).

نمونه‌گیری انباشته‌ای

تعیین عناصر تحقیق در گروه‌های خوشه‌ای مانند شهر، بخش، ناحیه و غیره غالباً از طریق روش انباشته‌ای صورت می‌گیرد، زیرا در این روش انتخاب تصادفی به‌طور خودکار متناسب با بزرگی هر گروه، انجام می‌شود.

به نحوه روش نمونه‌گیری می‌توان از طریق ذکر یک مثال بهتر پی برد. قبلاً باید به این موضوع اشاره کرد که با تعیین بزرگی جامعه نمونه تعداد مکانهای جغرافیایی را که در آن پرسشگری صورت می‌گیرد، نیز باید مشخص شود، زیرا تعداد محللهایی که در آن پرسشگری صورت می‌گیرد، به تعداد پرسشگر بستگی دارد.

اگر تعداد جامعه نمونه به عنوان مثال ۲۰۰۰ نفر و تعداد پرسشگر ۲۰۰ نفر باشد، در این صورت می‌توان در ۲۰۰ محل هر بار ۱۰ مصاحبه انجام داد، و اگر فقط ۱۰۰ پرسشگر در اختیار باشد، در این صورت می‌توان در ۱۰۰ محل هر بار ۲۰ مصاحبه انجام داد. اگر تعداد پرسشگر ۱۰۰ نفر باشد و ۲۰۰ محل انتخاب بشود، در این حالت هر پرسشگر باید در دو محل پرسشگری کند و این موجب افزایش هزینه مسافرت خواهد بود.

اگر مکان تحقیق چندین شهرستان باشد باید از بین این شهرستانها تعدادی نمونه انتخاب کرد (مثلاً تعدادی خانوار). برای این منظور تعداد خانوارهای هر یک از شهرستانها را مرحله به مرحله با هم جمع می‌کنیم. مجموع هر دو عدد (= مجموع قبلی + شهرستان آخری) را در ستون دیگری یادداشت می‌کنیم. این روش انباشته کردن، نامیده می‌شود. در اینجا خانوارهای شهرستانها به‌طور فرضی شماره‌گذاری می‌شوند. آخرین مجموع، تعداد خانوارهایی را نشان می‌دهد که در کلیه شهرستانهای مورد تحقیق زندگی می‌کنند. به عبارت دیگر مجموع نهایی تعداد جامعه اصلی N را نشان می‌دهد که از آن n خانوار انتخاب می‌شود. تعداد مصاحبه‌ای که در هر محل باید انجام بشود معلوم می‌کند که چند شهرستان را باید انتخاب کرد. نسبت بین N (جامعه

اصلی) و تعداد مصاحبه هر محل $N - K (m)$ شاخصی است که فاصله بین در واحد نمونه را مانند نمونه گیری منظم، مشخص می کند. در اینجا نیز عدد اول به طور تصادفی انتخاب می شود، که کوچکتر یا برابر با K است. عدد انتخابی را به K اضافه می کنیم و در نتیجه محل اولین مصاحبه بدست می آید. مجموع آخر را مجدداً با K جمع می کنیم. محل دومین مصاحبه را تعیین کنیم. این عمل را آنگذر ادامه می دهیم تا بالاخره مجموعی که از این طریق بدست می آید بزرگتر از تعداد جامعه اصلی N بشود. چون جمعیت شهرستانها متفاوت است، لذا ممکن است که یک شهرستان در نمونه گیری انتخاب نشود و برعکس شهرستان دیگری چندین بار انتخاب شود.

به عنوان مثال از ۱۰ شهرستان و با ۲۰ پرسشگر می خواهیم با ۲۰ خانوار مصاحبه کنیم. با بکار گرفتن روش انباشته ای به طریق زیر عمل می کنیم:

شهرستان	جمعیت	جمعیت انباشته
۱	۱۰۹۵۳	۱۰۹۵۳
۲	۱۹۷۶	۱۲۹۲۹
۳	۹۶۲۶	۲۲۵۵۵
۴	۴۰۶۶	۲۶۶۱۹
۵	۲۲۰۲	۲۸۸۲۱
۶	۱۳۹۳۴	۴۲۷۵۵
۷	۵۹۵۴	۴۹۷۰۹
۸	۱۳۶۶	۵۱۰۷۵
۹	۲۰۸۳۸	۷۱۹۱۳
۱۰	۳۷۳۴	۷۵۶۴۶

شاخصی فاصله بین واحدهای انتخابی در این مثال $۳۷۸۲ = ۲۰ \div ۷۵۶۴۶$ می باشد. اولین شماره را که باید بین ۱ و ۳۷۸۲ باشد، به طور تصادفی انتخاب می کنیم. مثلاً ۳۸۷. به ۳۸۷ عدد ۳۷۸۲ را می افزایم و کار را ادامه می دهیم، که در نتیجه اعداد ۳۸۷، ۴۱۶۹، ۷۹۵۱، ۱۱۷۳۳ و غیره بدست می آید. اعداد ۳۸۷، ۴۱۶۹، ۷۹۵۱، ۱۱۷۳۳ اول قرار دارند. عدد ۱۱۷۳۳ در شهرستان دوم، و الی آخر. از اینرو در شهرستان اول در ۳ محل مصاحبه، یعنی $۳ \times ۲۰ = ۶۰$ پرسشگری صورت می گیرد. بدیهی است که روش انباشته ای را می توان در مورد جامعه قشر بندی شده نیز بکار گرفت. تعداد نمونه ها در هر شهرستان از طریق قشر بندی منطقه ای تعیین می شود، و در نتیجه تعداد مصاحبه در هر منطقه مشخص می گردد. تعداد نمونه شهرستانهای هر منطقه از

طریق روش انباشته‌ای قابل تعیین است. نمونه‌هایی که در شهرستانها بدون قشر بندی تعیین می‌شود، معمولاً با روش نمونه‌گیری قشر بندی شده تفاوت زیادی ندارد. واریانس کل درحالتی که در داخل گروهها مجدداً نمونه‌گیری صورت نگیرد، از طریق فرمول زیر قابل محاسبه است (۵):

$$S_{\bar{X}} = \sqrt{\left(1 - \frac{m}{M}\right) \frac{\sigma_b}{m}}$$

در این فرمول M تعداد گروههای خوشه‌ای، m تعداد گروههای انتخاب شده و σ_b میانگین پراکندگی گروهها است.

$$\sigma_b^2 = \frac{1}{m-1} \sum_{i=1}^m (\bar{X}_i - \bar{X})^2$$

محاسبه واریانس کل، در صورتیکه در هر گروه مجدداً نمونه‌گیری دیگری صورت گیرد از طریق فرمول زیر بدست می‌آید:

$$S_{\bar{X}} = \sqrt{\left(1 - \frac{m}{M}\right) \frac{\sigma_h^2}{m} + \left(1 - \frac{n}{N}\right) \frac{\sigma_z^2}{m n}}$$

بنابراین واریانس کل با مجموع واریانس گروهها و واریانس جامعه نمونه (n) برابر است (۵).

تعداد نمونه‌ها

در پژوهشهای مبتنی بر نمونه‌گیری گویا، تعداد نمونه‌ها از مسائلی اساسی است. تعداد جامعه نمونه، تابع درجه اشتباه، درجه اطمینان، واریانس جامعه اصلی و تعداد جامعه اصلی، می‌باشد. بنابراین قبل از تعیین تعداد نمونه‌ها، باید واریانس جامعه اصلی را محاسبه کرد:

$$s^2 = \frac{\sum X^2 - (\sum X)^2}{N-1}$$

تعداد نمونه‌ها براساس فرمول زیر تعیین می‌شود:

$$n = \frac{t^2 \cdot N \cdot s^2}{t^2 \cdot s^2 + (N-1) \cdot e^2}$$

در این فرمول t درجه اطمینان و c درجه اشتباه است. بزرگی جامعه نمونه می تواند شامل تمام جامعه اصلی بشود، زیرا درجه گویایی نتایج جامعه نمونه در درجه اول تابع بزرگی جامعه نمونه است، و نه تابع نسبت بین جامعه اصلی و جامعه نمونه. بنا بر این انتخاب n از بین یک جامعه پنجاه هزار و یا صد هزار نفری تفاوتی ندارد. در حالتی که واریانس جامعه اصلی زیاد باشد، نمونه گیری از طریق روش قشر بندی شده صورت می گیرد. در این صورت و به طور کلی هنگامی که جامعه اصلی قبل از نمونه گیری گروه بندی می شود، واریانس جامعه اصلی که مجموعی از واریانس درونی و واریانس میانی قشرها است تغییر می کند.

محاسبه واریانس جامعه اصلی قشر بندی شده از طریق فرمول زیر بدست می آید:

$$s^2 = \frac{1}{N-1} \left\{ \sum N_i X_i^2 - \frac{1}{n} (\sum N_i X_i)^2 \right\}$$

در این فرمول N_i تعداد عناصر هر گروه یا قشر و X_i ارزش متوسط هر گروه است.

درجه اشتباه و درجه اطمینان را میتوان بر اساس منحنی نرمال به صورت جدول زیر نشان داد.

درجه اشتباه: (c)	درجه اطمینان به صورت درصد سطح منحنی نرمال	درجه اطمینان (t)
$\pm 3 \quad 7$	۶۸/۳	۱
± 5	۹۵	۱/۹۶
$\pm 4/5$	۹۵/۵	۲
± 1	۹۹	۲/۵۷
$\pm 0/۳$	۹۹/۷	۳
$\pm 0/1$	۹۹/۹	۳۲/۹

با داشتن واریانس، درجه اطمینان، میانگین ارزشها و تعداد نمونه حد اعتماد از طریق فرمول زیر قابل محاسبه است:

$$\text{حد اعتماد} = \bar{x} \pm t \frac{s}{\sqrt{n}}$$

در صورتیکه در نمونه‌گیری تصادفی تعداد نمونه‌ها بیش از ده درصد جامعه اصلی (یا قشرها) باشد، تعداد عناصر نمونه را میتوان براساس عامل اصلاح $\frac{1}{1 + \frac{n}{N}}$ ، کاهش داد (۹). در این صورت تعداد نهایی نمونه‌ها با فرمول زیر بدست می‌آید:

$$n' = n \left(\frac{1}{1 + \frac{n}{N}} \right)$$

$$n' = \frac{n}{1 + \frac{n}{N}}$$

زیر نویس‌ها:

- 1- W. Etzel, W. Joenk, M.D.U. Naeve, P. Statistische Tabellen. Berlin. 1967.
- 2 - Hampe, A. Statistik fuer Betriebswirte II. Stuttgart. 1974.
- 3 - Kellerer, H. Statistik im modernen Wirtschafts und Sozialleben. Hamburg. 1976.
- 4- René Koenig. Handbuch der empirischen Sozial forschung Band. 3A. Stut.gart. 1974.
- 5 - Holm, K. Die Befragung. 1. Muenchen. 1975.
- 6 - Stange, K. Angewandte Statistik zweiter Teil. Berlin. New York. 1971.
- 7 - Noelle, E. Umfrage jn der Massengesellschaft. Hamburg. 1973.
- 8 - Hill, W. Marketing. I. Stuttgart. 1973.
- 9 - W. Allen, Wallis Harry, V. Roberts. Methoden der Statistik. Stuttgart. 1972.