



دانشگاه سوادکوه

تحلیل‌های آماری در برنامه‌ریزی روستایی

(کارشناسی ارشد جغرافیا)

دکتر مصطفی طالشی دکتر پرویز نصیری

امروزه کتاب‌خوانی و علم‌آموزی، نه تنها یک وظیفه‌ی ملی، که یک واجب دینی است.

مقام معظم رهبری

در عصر حاضر یکی از شاخصه‌های ارزیابی رشد، توسعه و پیشرفت فرهنگی هر کشوری میزان تولید کتاب، مطالعه و کتاب‌خوانی مردم آن مرز و بوم است. ایران اسلامی نیز از دیرباز تاکنون با داشتن تمدنی چندهزارساله و مراکز متعدد علمی، فرهنگی، کتابخانه‌های معتبر، علما و دانشمندان بزرگ با آثار ارزشمند تاریخی، سرآمد دولت‌ها و ملت‌های دیگر بوده و در عرصه‌ی فرهنگ و تمدن جهانی به‌سان خورشیدی تابناک همچنان می‌درخشد و با فرزندان نیک‌نهاد خویش هنرنمایی می‌کند. چه کسی است که در دنیا با دانشمندان فرزانه و نام‌آور ایرانی همچون ابوعلی سینا، ابوریحان بیرونی، فارابی، خوارزمی و ... همچنین شاعران برجسته‌ای نظیر فردوسی، سعدی، مولوی، حافظ و ... آشنا نباشد و در مقابل عظمت آنها سر تعظیم فرود نیاورد. تمامی این افتخارات ارزشمند، برگرفته از میزان عشق و علاقه فراوان ملت ما به فراگیری علم و دانش از طریق خواندن و مطالعه منابع و کتاب‌های گوناگون است. به شکرانه‌ی الهی، تاریخ و گذشته ما، همیشه درخشان و پر بار است. ولی اکنون در این زمینه در چه جایگاهی قرار داریم؟ آمار و ارقام ارائه‌شده از سوی مجامع و سازمان‌های فرهنگی در مورد سرانه‌ی مطالعه‌ی هر ایرانی، برایمان چندان امیدوارکننده نمی‌باشد و رهبر معظم انقلاب اسلامی نیز از این وضعیت بارها اظهار گله و ناخشنودی نموده‌اند.

کتاب، دروازه‌ای به سوی گستره‌ی دانش و معرفت است و کتاب خوب، یکی از بهترین ابزارهای کمال بشری است. همه‌ی دستاوردهای بشر در سراسر عمر جهان، تا آنجا که قابل کتابت بوده است، در میان دست‌نوشته‌هایی است که انسان‌ها پدید آورده و می‌آورند. در این مجموعه‌ی بی‌نظیر، تعالیم الهی، درس‌های پیامبران به بشر، و همچنین علوم مختلفی است که سعادت بشر بدون آگاهی از آنها امکان‌پذیر نیست. کسی که با دنیای زیبا و زندگی‌بخش کتاب ارتباط ندارد بی‌شک از مهم‌ترین دستاورد انسانی و نیز از بیشترین معارف الهی و بشری محروم است. با این دیدگاه، به‌روشنی می‌توان ارزش و مفهوم رمزی عمیق در این حقیقت تاریخی را دریافت که اولین خطاب خداوند متعال به پیامبر گرامی اسلام (ص) این است که «بخوان!» و در اولین

سوره‌ای که بر آن فرستاده‌ی عظیم‌الشان خداوند، فرود آمده، نام «قلم» به تجلیل یاد شده‌است: «اقْرَأْ وَ رَبُّكَ الْأَكْرَمُ. الَّذِي عَلَّمَ بِالْقَلَمِ» در اهمیت عنصر کتاب برای تکامل جامعه‌ی انسانی، همین بس که تمامی ادیان آسمانی و رجال بزرگ تاریخ بشری، از طریق کتاب جاودانه مانده‌اند.

دانشگاه پیام‌نور با گستره‌ی جغرافیایی ایران‌شمول خود با هدف آموزش برای همه، همه‌جا و همه‌وقت، به‌عنوان دانشگاهی کتاب‌محور در نظام آموزش عالی کشورمان، افتخار دارد جایگاه اندیشه‌سازی و خردورزی بخش عظیمی از جوانان جویای علم این مرز و بوم باشد. تلاش فراوانی در ایام طولانی فعالیت این دانشگاه انجام پذیرفته تا با بهره‌گیری از تجربه‌های گرانقدر استادان و صاحب‌نظران برجسته کشورمان، کتاب‌ها و منابع آموزشی درسی شاخص و خودآموز تولید شود. در آینده هم، این مهم با هدف ارتقای سطح علمی، روزآمدی و توجه بیشتر به نیازهای مخاطبان دانشگاه پیام‌نور با جدیت ادامه خواهد داشت. به‌طور قطع استفاده از نظرات استادان، صاحب‌نظران و دانشجویان محترم، ما را در انجام این وظیفه‌ی مهم و خطیر یاری‌رسان خواهد بود. پیشاپیش از تمامی عزیزانی که با نقد، تصحیح و پیشنهادهای خود ما را در انجام این وظیفه‌ی خطیر یاری می‌رسانند، سپاسگزاری می‌نماییم. لازم است از تمامی اندیشمندانی که تاکنون دانشگاه پیام‌نور را منزلگاه اندیشه‌سازی خود دانسته و ما را در تولید کتاب و محتوای آموزشی درسی یاری نموده‌اند، صمیمانه قدردانی گردد. موفقیت و بهروزی تمامی دانشجویان و دانش‌پژوهان عزیز آرزوی همیشگی ما است.

دانشگاه پیام‌نور

فهرست مطالب

پیشگفتار	نه
هدف کلی و هدف‌های یادگیری درس	یازده
فصل اول. آمار توصیفی	۱
هدف کلی	۱
هدف‌های یادگیری	۱
مقدمه	۱
۱-۱ مفاهیم اساسی	۲
۱-۱-۱ جامعه آماری	۲
۱-۱-۲ نمونه	۳
۱-۱-۳ داده‌های آماری	۳
۱-۱-۴ متغیر	۴
۲-۱ نمونه‌گیری تصادفی ساده	۵
۳-۱ نمونه‌گیری با جای‌گذاری و بدون جای‌گذاری	۵
۴-۱ انتخاب نمونه تصادفی ساده	۵
۵-۱ نمونه‌گیری منظم یا سیستماتیک	۷
۶-۱ نمونه‌گیری تصادفی با طبقه‌بندی	۸
۷-۱ نمونه‌گیری تصادفی خوشه‌ای	۱۰
۸-۱ شاخص‌های گرایش مرکز	۱۱
۱-۸-۱ میانگین حسابی	۱۱
۲-۸-۱ میانگین هندسی	۱۲
۳-۸-۱ میانگین هارمونیک	۱۳
۴-۸-۱ نما	۱۵
۵-۸-۱ میانه	۱۵

۱۶	 ۱-۸-۶ چارک‌ها
۱۸	 ۱-۹-۹ شاخص‌های پراکندگی
۱۸	 ۱-۹-۱ دامنه
۱۸	 ۱-۹-۲ واریانس نمونه
۲۰	 ۱-۹-۳ انحراف معیار
۲۱	 ۱-۹-۴ ضریب تغییرات
۲۲	 ۱-۱۰-۱ جدول توزیع فراوانی
۲۷	 ۱-۱۱-۱ محاسبه میانگین و واریانس در جدول توزیع فراوانی
۲۹	 ۱-۱۲-۱ محاسبه نما در جدول توزیع فراوانی
۳۰	 ۱-۱۳-۱ محاسبه میانه، چارک‌ها و صدک‌ها در جدول توزیع فراوانی
۳۲	 ۱-۱۴-۱ نمودارهای توصیفی
۳۲	 ۱-۱۴-۱ نمودار میله‌ای (ستونی)
۳۳	 ۱-۱۴-۲ نمودار دایره‌ای
۳۵	 ۱-۱۴-۳ نمودار مستطیلی
۳۶	 ۱-۱۴-۴ نمودار چندضلعی فراوانی
۳۸	 ۱-۱۴-۵ نمودار فراوانی تراکمی (اجایو)
۴۱	 خودآزمایی فصل اول
۴۷	 پیوست نرم‌افزار SPSS

۵۳	 فصل دوم. آزمون فرضیه آماری
۵۳	 هدف کلی
۵۳	 هدف‌های یادگیری
۵۳	 مقدمه
۵۴	 ۱-۲ مفاهیم اولیه آزمون فرض
۵۶	 ۲-۲ آزمون فرض درباره میانگین جامعه نرمال
۶۱	 ۲-۳ آزمون فرض برای برابری میانگین‌های دو جامعه نرمال
۶۵	 ۲-۳ آزمون فرض درباره نسبت جامعه و مقایسه نسبت‌های دو جامعه
۶۹	 ۲-۴ آزمون فرض درباره واریانس جامعه نرمال و مقایسه واریانس‌های دو جامعه
۷۳	 خودآزمایی فصل دوم
۷۹	 پیوست نرم‌افزار SPSS

۸۹	 فصل سوم. روش‌های ناپارامتری
۸۹	 هدف کلی
۸۹	 هدف‌های یادگیری
۹۰	 مقدمه
۹۱	 ۱-۳ آزمون علامت
۹۴	 ۳-۲ آزمون رتبه-علامت ویلکاکسون
۹۷	 ۳-۳ آزمون ویلکاکسون یا آزمون من-ویتنی

۳-۴	آزمون کولموگروف- اسمیرنف	۱۰۰
۳-۵	آزمون کراسکال- والیس یا فریدمن	۱۰۱
۳-۶	آزمون ککران	۱۰۳
۱۰۷	خودآزمایی فصل سوم	
۱۰۹	پیوست نرم افزار SPSS	

۱۱۷	فصل چهارم. آنالیز واریانس	
۱۱۷	هدف کلی	
۱۱۷	هدف های یادگیری	
۱۱۸	مقدمه	
۱۱۹	۱-۴ آنالیز واریانس یک طرفه	
۱۲۰	۲-۴ مجموع مربعات و جدول ANOVA	
۱۲۵	۳-۴ آزمون دانکن برای تفاوت های معنی داری بین زوج میانگین ها	
۱۳۱	۴-۴ تحلیل واریانس برای داده های رتبه بندی شده	
۱۳۵	خودآزمایی فصل چهارم	
۱۳۹	پیوست نرم افزار SPSS	

۱۴۵	فصل پنجم. همبستگی	
۱۴۵	هدف کلی	
۱۴۵	هدف های یادگیری	
۱۴۶	مقدمه	
۱۴۷	۱-۵ نمودار پراکنش	
۱۴۸	۲-۵ ضریب همبستگی نمونه ای پیرسن	
۱۵۰	۳-۵ آزمون فرض درباره ضریب همبستگی جامعه	
۱۵۲	۳-۵ رتبه گذاری	
۱۵۳	۴-۵ ضریب همبستگی رتبه ای اسپیرمن	
۱۵۴	۴-۵ آزمون فرض درباره ضریب همبستگی رتبه ای اسپیرمن در جامعه	
۱۵۵	۵-۵ ضریب همبستگی رتبه ای کندال	
۱۵۷	۵-۵ آزمون فرض درباره ضریب همبستگی رتبه ای کندال در جامعه	
۱۵۹	۶-۵ ضریب همبستگی جزئی	
۱۶۱	خودآزمایی فصل پنجم	
۱۶۴	پیوست نرم افزار SPSS	

۱۶۹	فصل ششم. رگرسیون	
۱۶۹	هدف کلی	
۱۶۹	هدف های یادگیری	
۱۶۹	مقدمه	
۱۷۰	۱-۶ معادله خط برازش	

۱۷۶	۲-۶ آزمون فرض درباره β_1
۱۷۷	۳-۶ آنالیز رگرسیون
۱۸۰	۴-۶ مدل رگرسیون چندمتغیره
۱۸۲	خودآزمایی فصل ششم
۱۸۵	پیوست نرم افزار SPSS
۱۹۳	فصل هفتم. روش تحلیل عاملی
۱۹۳	هدف کلی
۱۹۳	هدف های یادگیری
۱۹۴	مقدمه
۱۹۴	۱-۷ ماتریس کوواریانس - واریانس و ماتریس ضریب همبستگی
۱۹۶	۲-۷ ماتریس کوواریانس - واریانس نمونه و ماتریس ضریب همبستگی
۲۰۰	۳-۷ مقادیر و بردار ویژه یا خاص ماتریس
۲۰۲	۴-۷ الگوی عاملی
۲۰۵	۵-۷ تعیین تعداد عامل ها
۲۰۵	۱-۵-۷ ضریب KMO
۲۰۶	۲-۵-۷ چرخش عاملی
۲۰۹	خودآزمایی فصل هفتم
۲۱۳	پیوست نرم افزار SPSS
۲۱۹	فصل هشتم. تحلیل خوشه ای
۲۱۹	هدف کلی
۲۱۹	هدف های یادگیری
۲۲۰	مقدمه
۲۲۰	۱-۸ ماتریس داده
۲۲۱	۲-۸ کمی سازی ماتریس داده یا تصمیم گیری
۲۲۳	۳-۸ استاندارد سازی ماتریس داده یا تصمیم گیری
۲۲۵	۴-۸ آنالیز خوشه ای
۲۲۹	۵-۸ کاربرد نرم افزار در آنالیز خوشه ای
۲۳۰	خودآزمایی فصل هشتم
۲۳۳	پیوست ۱. آشنایی با نرم افزار SPSS
۲۳۴	نحوه نصب نرم افزار
۲۳۵	معرفی نوار بالای نرم افزار
۲۴۹	پیوست ۲. جداول آماری
۲۶۵	منابع

پیشگفتار

در بازشناسی به‌روش‌شناسی تحقیق در علوم جغرافیایی، همواره دو الگوی مطالعات و پژوهش‌های کمی و کیفی، قابل طبقه‌بندی است. البته گفتنی است در بسیاری از مباحث و تحقیقات جغرافیایی، به‌لحاظ ماهیت‌شناسی این دو الگو به‌شیوه ترکیبی نیز مورد بهره‌برداری قرار می‌گیرند.

در روش مطالعات کمی به‌لحاظ محتوا و ارزش‌گذاری مطلوب برای دستیابی به‌نتایج پژوهش قابل اعتبار، نقش مطالعات و تحلیل‌های آماری از اهمیت فراوانی برخوردار است. از این رو شناخت روش‌ها و ابزارهای آمار توصیفی و تحلیلی به‌ویژه با کاربرد پژوهشی آن‌ها در مطالعات و برنامه‌ریزی روستایی از دیدگاه علوم جغرافیایی بسیار حائز اهمیت است. البته نکته اساسی در شناخت این الگوها و روش‌های آماری، آن است که نبایستی در حصار و دیوارهای بلند تحلیل‌های کمی صرف، جغرافی‌دان خود را محدود کند، زیرا با بهره‌گیری از تحلیل‌های آماری و با شناخت دقیق و کیفی مباحث در نواحی جغرافیایی، بنیان‌های دقیق تحلیلی آن موضوع در برنامه‌ریزی روستایی مورد ارزیابی قرار می‌گیرد.

با این پیش‌زمینه، این گفتار با عنوان تحلیل‌های آماری در برنامه‌ریزی روستایی سازمان‌دهی شده‌است. در این زمینه، تلاش بر آن است که متخصصان جغرافیا و برنامه‌ریزی روستایی بتوانند با شناخت روش‌های آماری مورد نیاز، تحلیل‌های کاربردی خود را صورت‌بندی عملیاتی‌کنند تا به‌وسیله تحلیل و موضوع‌شناسی دقیق، مباحث برنامه‌ریزی روستایی و پیش‌بینی تحقیق سازمان‌بخشی توسعه پایدار روستایی را سازمان‌دهی‌کنند.

از سوی دیگر در این گفتار، هدف آن بوده است که بخش قابل توجهی از موضوعات اساسی در تحلیل آماری مورد نیاز در برنامه ریزی روستایی آورده شود. البته گفتنی است با توجه به پیچیدگی و گستردگی تحلیل های آماری پیشرفته، همواره تلاش شده است با بهره گیری از ابزارهای نرم افزاری مناسب، بخش قابل توجهی از نیازهای تحقیقاتی دانشجویان کارشناسی ارشد در این زمینه فراهم شود.

به همین منظور، مباحث در هشت فصل تهیه و تنظیم شده است. در فصل اول، به آمار توصیفی پرداخته شده است. فصل دوم، آزمون فرض های آماری را مورد توجه قرار می دهد. در ادامه، در فصل سوم و چهارم به ترتیب روش های ناپارامتری و آنالیز واریانس مورد بازشناسی قرار می گیرد. در فصل پنجم و ششم نیز مباحث اساسی همبستگی و رگرسیون، مورد بررسی قرار می گیرد. فصول انتهایی کتاب نیز مربوط به تحلیل عاملی و خوشه ای است.

در اینجا نگارندگان وظیفه خود می دانند که از زحمات بی شائبه همکاران ارجمند خود در مدیریت تدوین و تولید کتب و محتوای آموزشی دانشگاه پیام نور که با کمک و مساعدت ایشان، امکان تحریر این گفتار فراهم آمد، قدردانی کنند. همچنین از سرکار خانم سارا طالشی، دانش آموخته آمار دانشگاه شهید بهشتی و خانم فاطمه محمدی که حروف چینی و صفحه آرایی کتاب را برعهده داشتند تشکر کنند و همچنین جا دارد از زحمات ارزنده جناب آقای حاج سلمانی در تنظیم برنامه نرم افزاری SPSS قدردانی شود. در پایان با فرض پذیرش کاستی های اولیه، امید است صاحب نظران دانشگاهی، سایر اندیشمندان علوم آماری، جغرافی دانان حوزه برنامه ریزی روستایی و کارشناسان اجرایی دست اندر برنامه ریزی روستایی، با راهنمایی های ارزشمند خود ما را در چاپ بعدی یاری کنند. پیشاپیش از بذل محبت همه همکاران و کارشناسان اجرایی دستگاه های پژوهشی دولتی و غیردولتی، بسیار سپاسگزاریم و پذیرای نظرات و پیشنهادهای این عزیزان به واسطه آدرس رایانامه زیر خواهیم بود.

مصطفی طالشی

Mtaleshi@pnu.ac.ir
taleshimo@yahoo.com

پرویز نصیری

pnasiri@pnu.ac.ir
pnasiri45@yahoo.com

هدف کلی و هدف‌های یادگیری درس

شناخت تحلیل آماری به‌ویژه در برنامه‌ریزی روستایی، یکی از ابزارهای اساسی در تحلیل کمی مطالعات و پژوهش‌های برنامه‌ریزی و توسعه روستایی به‌شمار می‌رود. هدف از طرح درس تحلیل‌های آماری در برنامه‌ریزی روستایی در چهارچوب مصوبه رشته جغرافیا و برنامه‌ریزی روستایی در مقطع کارشناسی ارشد در تمام گرایش‌های چهارگانه، آشنایی با روش‌های آماری در برنامه‌ریزی روستایی به‌ویژه کاربرد روش‌های آماری مبتنی بر آزمون فرضیه در تحقیقات روستایی است، زیرا شناخت و تسلط بر روش‌های آماری، از دیدگاه آمار توصیفی و آمار تحلیلی، توانمندی جغرافی‌دان روستایی را در تحلیل‌های جغرافیایی، بیش‌ازپیش تقویت می‌کند.

دستیابی به‌هدف یادشده، با استفاده از هدف‌های مرحله‌ای زیر قابل دسترس است:

۱. شناخت و بهره‌گیری از آمار توصیفی در شناخت مباحث مرتبط با جغرافیای روستایی و برنامه‌ریزی روستایی.
۲. بازشناسی آزمون‌های مرتبط با فرض‌های آماری در پژوهش‌های برنامه‌ریزی روستایی.
۳. شناخت روش ناپارامتری در تحقیقات و مطالعات برنامه‌ریزی روستایی.
۴. کاربردهای اساسی آنالیز واریانس در پژوهش‌های برنامه‌ریزی روستایی.
۵. کاربرد روش‌های همبستگی و رگرسیون در برنامه‌ریزی روستایی.
۶. کاربرد تحلیل‌های عاملی و خوشه‌ای در برنامه‌ریزی روستایی.

فصل اول

آمار توصیفی

هدف کلی

هدف اصلی از این فصل، بیان و شناخت روش‌ها و الگوهای آمار توصیفی است. از سوی دیگر در این فصل مفاهیمی اساسی همچون جامعه آماری، نمونه، داده‌های آماری و متغیر، مورد بررسی قرار می‌گیرد. روش‌های نمونه‌گیری در جامعه آماری به‌ویژه شاخص‌های گرایش مرکزی، شاخص‌های پراکندگی و انواع نمودارهای مورد استفاده در آمار توصیفی، در این فصل مورد توجه قرار می‌گیرد.

هدف‌های یادگیری

دانشجویان پس از مطالعه این فصل قادر خواهند بود به‌موارد ذیل پاسخ دهند: تعاریف و مفاهیم اساسی در آمار توصیفی را بیان کنند.

۱. انواع روش‌های نمونه‌گیری و کاربرد هریک را توضیح دهند.
۲. ویژگی‌ها و روش‌های شاخص‌های گرایش مرکز را توضیح دهند.
۳. شاخص‌های پراکندگی را در مطالعات و پژوهش‌های برنامه‌ریزی روستایی بیان کنند.
۴. روش‌های استفاده از انواع نمودارهای موضوعی را توضیح دهند.

مقدمه

آمار علمی است که با جمع‌آوری، تنظیم، تجزیه و تحلیل و تفسیر مشاهدات سروکار دارد. جمع‌آوری داده‌ها، مرحله تهیه اندازه‌گیری یا شمارش‌ها به‌شمار می‌رود. نتایج

قابل اطمینان، فقط از داده‌هایی که به‌طور صحیح جمع‌آوری شده‌اند، به‌دست می‌آیند. جمع‌آوری داده‌ها قسمت مهمی از روش آماری است که به‌وسیله نمونه‌گیری از جامعه آماری انجام می‌شود. مجموعه روش‌ها و قوانینی که نتایج را ساده‌تر کند و به‌تفسیر آسان‌تری از جمع‌آوری و طبقه‌بندی داده‌ها یا مشاهدات منجر شود، آمار توصیفی نامیده می‌شود. در به‌کارگیری علم آمار، این نکته حائز اهمیت است که هیچ روش آماری، به‌تنهایی از بروز اشتباهات، نادرستی‌ها، استدلال غلط یا نتایج غیرصحیح جلوگیری نمی‌کند؛ بنابراین بایستی مشاهدات آماری، به‌گونه‌ای انجام شود که ضریب خطای آماری را به‌حداقل ممکن کاهش یابد. از سوی دیگر، روش‌های آماری نیز به‌طور صحیح به‌کار برده شود و نتایج به‌وسیله کسی تجزیه و تحلیل و تفسیر شود که نه تنها با روش‌های آماری بلکه با رشته‌ای که آمار در مورد آن به‌کار رفته است آشنا باشد. در این فصل، مفاهیم اساسی، روش‌های نمونه‌گیری، شاخص‌های مرکزی و غیرمرکزی، جدول توزیع فراوانی و انواع نمودارهای آماری ارائه می‌شود.

۱-۱ مفاهیم اساسی

در این بخش برای آشنایی بیشتر با ادبیات آماری، اصطلاحات اساسی آمار که همواره مورد استفاده است بررسی می‌شوند.

۱-۱-۱ جامعه آماری

جامعه آماری، مجموعه یا کل اندازه‌گیری‌ها از تمام اشیائی است که حداقل یک صفت مشخص دارند. از نظر آمار، یک جامعه آماری هرگز مجموعه‌ای از اشیاء نیست، بلکه همیشه مجموعه‌ای از اندازه‌گیری‌ها یا شمارش‌هاست. یک جامعه آماری ممکن است از تعداد متناهی یا نامتناهی داده تشکیل شده باشد. اگر جامعه آماری، متناهی و دارای N عضو باشد اعضای آن را به‌صورت زیر نمایش می‌دهند:

$$X_1, X_2, \dots, X_N$$

که X_1 عضو اول، X_2 عضو دوم ... و X_N عضو N ام جامعه است و هیچ ترتیبی بین آن‌ها نیست.

مثال. فرض کنید انجمن برنامه‌ریزی روستایی کشور، دارای ۱۰ عضو باشد. در این مثال $N = 10$ و اعضای آن را می‌توان جداگانه به صورت زیر با نماد X_1 تا X_{10} نمایش داد:

$$X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}.$$

۱-۲ نمونه

نمونه، مجموعه‌ای از اندازه‌گیری‌هاست که قسمتی یا تمام یک جامعه را دربرمی‌گیرد. اگر نمونه انتخابی، قابل شمارش یا برچسب‌پذیر و دارای n ($n \leq N$) عضو باشد، اعضای نمونه را به صورت زیر نمایش می‌دهند:

$$X_1, X_2, \dots, X_n$$

مثال. اگر سه عضو از انجمن برنامه‌ریزی روستایی برای فعالیت در یک طرح پژوهشی مشترک، مشارکت داشته باشند، اعضای آن را می‌توان به صورت زیر نوشت:

$$n = 3, \quad X_1, X_2, X_3$$

۱-۳ داده‌های آماری

داده‌های آماری به دو گروه داده‌های دست‌اول یا خام و داده‌های دست‌دوم تقسیم‌بندی می‌شود. داده‌های دست‌اول، آن دسته از داده‌هایی هستند که روی آن‌ها هیچ‌گونه سازماندهی یا پردازش صورت نگرفته است. برای نمونه جمع‌آوری اطلاعات با استفاده از اطلاعات پرسشنامه‌ای به وسیله مرکز آمار ایران در سرشماری‌ها از هزینه زندگی خانوارهای روستایی، داده‌های دست‌اول یا خام و گزارش‌های چاپ‌شده به وسیله مرکز آماری ایران به صورت جداول یا نمودارها از نوع داده‌های دست‌دوم است.

جدول زیر که توزیع درآمد را بین خانوارهای روستایی استانی نشان می‌دهد به عنوان داده دست‌دوم تلقی می‌شود و می‌توان آن را با اعلام منبع، مورد استفاده قرار داد (ارقام به میلیون تومان).

درصد فراوانی	فراوانی	رده درآمدی
۱۰	۵۰	۵-۷
۲۶	۱۳۰	۷-۹
۳۴	۱۷۰	۹-۱۱
۳۰	۱۵۰	۱۱-۱۳
۱۰۰	۵۰۰	

۴-۱-۱ متغیر

متغیر صفتی است برای معرفی اعضای جامعه آماری و به‌ویژگی اطلاق می‌شود و تغییرات را از فردی به فرد یا از شیئی به شیء دیگر نشان می‌دهد. برای مثال میزان تولید یک محصول در واحد سطح یا میزان مهاجر فرستی سکونت‌گاه‌های روستایی، هریک به‌عنوان متغیر تلقی می‌شود. به‌دلیل تغییرات، این متغیر همواره به‌صورت کمی یا کیفی قابل بررسی است. متغیر کمی به‌متغیری گفته می‌شود که قابل اندازه‌گیری بوده یا برای اندازه‌گیری آن مقیاسی وجود داشته‌باشد.

متغیر کمی ممکن است گسسته یا پیوسته باشد. متغیر کمی گسسته متغیری است که قابل شمارش باشد؛ مانند میزان جمعیت روستایی، تعداد طرح‌های اجرایی روستایی، تعداد جمعیت بیکار و تعداد روزهای یخبندان.

متغیر کمی پیوسته، متغیری است که بتواند بین هر دو مقدار دلخواه، مقداری انتخاب‌کند؛ مانند وزن یک محصول، سرعت باد.

متغیر کیفی به‌متغیری اطلاق می‌شود که اختلاف و تغییرات بین مقادیر مختلف آن، کیفی است و برای ثبت آن ممکن است از روش‌های دیگری غیر از به‌کاربردن عدد استفاده‌شود، اما در نهایت به‌منظور تحلیل آماری، بایستی این متغیرهای کیفی نیز به‌مقادیر کمی تبدیل شوند. برای مثال میزان فرهنگ جوامع روستایی که باید به‌وسیله مقادیر کمی، مورد تحلیل قرارگیرد.

بدین ترتیب اگر متغیری با x نشان داده‌شود، مشاهدات متوالی این متغیر به‌صورت:

$$x_1, x_2, \dots, x_i, \dots, x_n$$

نشان داده می‌شود که x_i مشاهده i ام و x_n عضو یا مشاهده n ام و n حجم مشاهدات است.

قبل از بررسی شاخص‌های گرایش مرکزی و پراکندگی، روش‌های نمونه‌گیری که در این بخش، سهم مهمی از تحلیل آماری دارند، مورد بررسی قرار می‌گیرد. مهم‌ترین روش‌های نمونه‌گیری عبارت‌اند از:

۱. نمونه‌گیری تصادفی ساده^۱
۲. نمونه‌گیری منظم یا سیستماتیک^۲
۳. نمونه‌گیری تصادفی یا طبقه‌بندی^۳
۴. نمونه‌گیری تصادفی خوشه‌ای^۴

۱-۲ نمونه‌گیری تصادفی ساده

نمونه‌گیری تصادفی ساده، یکی از روش‌های متداول نمونه‌گیری است. در این روش فرض می‌شود هر عضو جامعه، شانس مساوی برای انتخاب شدن دارد. نمونه‌گیری تصادفی ساده ممکن است با جای‌گذاری یا بدون جای‌گذاری باشد.

۱-۳ نمونه‌گیری با جای‌گذاری و بدون جای‌گذاری

در نمونه‌گیری با جای‌گذاری، هر مشاهده i ام از جامعه پس از رؤیت یا ثبت، به‌جامعه برگردانده می‌شود. در این روش وجود مشاهدات تکراری، ممکن است. در نمونه‌گیری بدون جای‌گذاری پس از رؤیت یا ثبت نمونه i ام، عضو مشاهده‌شده از جامعه، به‌جامعه برگردانده نمی‌شود؛ بنابراین در نمونه‌گیری با جای‌گذاری، شانس انتخاب هر عضو، ثابت است؛ ولی در نمونه‌گیری بدون جای‌گذاری، ثابت نیست.

۱-۴ انتخاب نمونه تصادفی ساده

یکی از راه‌های انتخاب نمونه تصادفی ساده، استفاده از جدول اعداد تصادفی، ماشین حساب یا برنامه‌های آماده رایانه‌ای است. در این قسمت روش استفاده از جدول اعداد تصادفی (جدول شماره ۱) شرح داده می‌شود.

1. Simple Random Sampling
 2. Systematic Sampling
 3. Stratified Random Sampling
 4. Cluster Random Sampling

فرض کنید جامعه آماری مورد بررسی، شامل ۱۵ خانوار روستایی است که اطلاعات درآمد سرانه آن به‌روش پژوهش میدانی جمع‌آوری شده‌است. میزان درآمد سرانه نیز به‌صورت هزار تومان برحسب سرانه به‌صورت زیر است:

۱۵ ۱۴ ۱۳ ۱۲ ۱۱ ۱۰ ۹ ۸ ۷ ۶ ۵ ۴ ۳ ۲ ۱: شماره خانوار

۱۵۰ ۱۵۸ ۱۶۱ ۱۵۸ ۱۵۲ ۱۵۲ ۱۴۳ ۱۵۰ ۱۵۸ ۱۶۰ ۱۵۱ ۱۵۸ ۱۵۲ ۱۵۱ ۱۴۳: درآمد سرانه

می‌خواهیم نمونه‌ای تصادفی به‌حجم ۵ با استفاده از جدول اعداد تصادفی انتخاب کنیم. ابتدا لازم است در نقطه شروع تصادفی قرارگیریم. این مهم را به‌روش‌های زیادی می‌توان انجام داد. یکی از راه‌ها این است که با نوک مداد بر نقطه‌ای تصادفی در صفحه اعداد تصادفی اثرگذاریم. نقطه تصادفی، نزدیک‌ترین رقم به‌محلی است که مداد بر صفحه اثرمی‌گذارد. فرض کنید نقطه شروع تصادفی در تقاطع سطر ششم و ستون سوم باشد؛ رقم مزبور در این نقطه صفر است. از آنجا که نمونه باید از میان ۱۵ خانوار انتخاب شود، تنها می‌توان اعداد ۱ تا ۱۵ را از جدول اعداد تصادفی استخراج کرد.

بنابراین باید اعداد دورقمی را چنان انتخاب کرد که در محدوده ۱ تا ۱۵ قرارگیرند. نخستین عدد دورقمی که در نقطه شروع تصادفی اعداد جدول انتخاب می‌شود، عدد ۰۲ یا ۲ است. اعداد بعدی به‌ترتیب عبارت‌اند از:

۷۱	۰۱	۷۹	۹۷	۷۰	۲۲	۷۳	۵۵	۷۷	۰۲	۳۹	۹۴	۰۲
۱۷	۵۴	۴۸	۱۷	۴۵	۸۱	۸۰	۲۱	۸۰	۷۵	۵۲	۵۲	۱۹
۶۹	۲۹	۵۱	۳۳	۰۵	۴۳	۷۱	۳۳	۹۹	۸۰	۱۱	۵۶	۸۴
										۷۱	۱۲	۵۶

پس از انتخاب اولین عدد تصادفی، سایر اعداد با حرکت از نقطه شروع به‌سمت راست انتخاب می‌شوند. می‌توان جهت حرکت را هرطور که از قبل تعیین شده‌است، رو به‌بالا یا به‌پایین ادامه داد. نمونه انتخابی متناظر با اعداد فوق عبارت‌اند از: ۲، ۱، ۱۱، ۵، ۱۲

عدد تصادفی انتخاب شده	۲	۱	۱۱	۵	۱۲
مقدار نمونه	۱۵۱	۱۴۳	۱۵۲	۱۵۱	۱۵۸

یادآوری می‌شود که ۰۲ یا عدد ۲، دو بار تکرار شده‌است که یکی از آن‌ها حذف شد. بنابراین نمونه انتخابی، برابر است با:

۱۵۱ ۱۴۳ ۱۵۲ ۱۵۱ ۱۵۸

۱-۵ نمونه‌گیری منظم یا سیستماتیک

در این روش تمام اعضای جامعه بدون هرگونه نظم و ترتیبی، فهرست‌بندی و سپس نمونه مورد نظر با استفاده از نظم معینی از لیست جامعه انتخاب می‌شود. اگر اعضای جامعه در لیست یا فهرست‌بندی، به‌صورت زیر باشد:

$$X_1, X_2, \dots, X_N$$

و بخواهیم یک نمونه به‌حجم n انتخاب کنیم، ابتدا عدد K یا تناوب، از رابطه زیر محاسبه می‌شود:

$$K = \left[\frac{N}{n} \right]$$

پس از محاسبه K ، جامعه به K قسمت تقسیم و از هر دسته، یک نمونه انتخاب می‌شود تا نمونه‌ای به‌حجم n به‌دست‌آید. نقطه شروع، به‌تصادف از بین ۱ تا K انتخاب می‌شود. عضو جامعه متناظر در لیست، عضو اول جامعه خواهد بود. برای مثال، فرض کنید پژوهشگری قصد دارد نگرش روستائیان به حجم $N=1000$ نفر را نسبت به فعالیت‌های اجرایی جهاد کشاورزی، مورد پژوهش قرار دهد. برای انتخاب نمونه‌ای به‌حجم $n=50$ ، فهرستی از همه روستائیان، تهیه و آن‌گاه از هر $\frac{N}{n} = \frac{1000}{50} = 20$ نفر، یک نفر روستایی را به‌صورت تصادفی انتخاب می‌کند. برای جلوگیری از هرگونه جهت‌گیری، یک عدد به‌تصادف از ۱ تا ۲۰ انتخاب و متناظر با آن و با اضافه‌کردن ۲۰ به آن، نمونه‌های متناظر از جامعه یا لیست انتخاب می‌شود. اگر عدد تصادفی انتخاب‌شده ۰۷ باشد، اعداد بعدی عبارت‌اند از:

۰۷ ۲۷ ۴۷ ۶۷ ۸۷ ۱۰۷ ۱۲۷ ۱۴۷ ...

اعضای متناظر با اعداد گفته‌شده، یعنی $X_{۱۰۷}$ ، $X_{۸۷}$ ، $X_{۶۷}$ ، $X_{۴۷}$ ، $X_{۳۷}$ ، $X_{۱۷}$ و $X_{۱۴۷}$ ، $X_{۱۲۷}$... نمونه‌های مورد بررسی خواهند بود.

هنگام طرح‌ریزی برای انتخاب یک نمونه از یک فهرست یا امثال آن، پژوهشگر باید با دقت فهرست را بررسی‌کند و مطمئن شود که هیچ الگوی چرخشی در آن وجود ندارد. به عبارت دیگر، در صورتی که دارای یک ترتیب تناوبی باشد، نمونه‌گیری منظم، عملکرد و نتیجه مناسبی نخواهد داشت.

مثال. فرض کنید می‌خواهید از ۲۰ سکونت‌گاه روستایی دهستان الف، ۵ سکونت‌گاه را به تصادف به روش سیستماتیک انتخاب کنید.

۱- محمدآباد	۶- حمیدآباد	۱۱- شاه‌تقی	۱۶- فرهادگرد
۲- تقی‌آباد	۷- داوودی	۱۲- طاهرآباد	۱۷- قاسم‌آباد
۳- اسماعیل‌آباد	۸- کلاته علی‌آباد	۱۳- خلیل‌آباد	۱۸- کریم‌آباد
۴- خرو سفلی	۹- کلاته شاه‌رخ	۱۴- علی‌آباد	۱۹- نورآباد
۵- خرو علیا	۱۰- صابری	۱۵- منصوریه	۲۰- یوسف‌آباد

در این مثال $N = ۲۰$ ، $n = ۵$ و در نتیجه $K = \frac{۲۰}{۵} = ۴$. برای شروع انتخاب، یک عدد بین ۱ و ۴ به تصادف انتخاب می‌شود. فرض کنید عدد انتخاب‌شده ۳ باشد؛ بنابراین اولین عضو نمونه، روستایی است که در چهارچوب شماره آن ۳ باشد که در این مثال، اسماعیل‌آباد اولین روستایی است که انتخاب‌شده و می‌تواند اطلاعات جمعیتی، اشتغال و غیره را که مدنظر پژوهشگران است از این سکونت‌گاه پرسشگری کرد. برای انتخاب سکونت‌گاه بعدی کافی است به شماره ۳، پیوسته عدد ۴ اضافه شود تا شماره سکونت‌گاه به دست‌آید.

۱۹ ۱۵ ۱۱ ۷ ۳

با توجه به شماره، سکونت‌گاه‌های انتخاب‌شده عبارت‌اند از: اسماعیل‌آباد، داوودی، شاه‌تقی، منصوریه و نورآباد.

۱-۶ نمونه‌گیری تصادفی با طبقه‌بندی^۱

در این روش، جامعه آماری از لحاظ تشابه یا براساس یک یا چند صفت متغیر،

طبقه‌بندی و نمونه‌های مستقل، از هر طبقه انتخاب می‌شود. در این روش، جامعه آماری به حجم N ، به h طبقه یا زیرجامعه آماری همگن به حجم‌های $N_1, N_2, \dots, N_h$ افراز و از هر طبقه، نمونه‌ای متناسب با حجم زیرجامعه انتخاب می‌شود.

اگر $n_1, n_2, \dots, n_h$ به ترتیب نمونه‌های انتخابی از زیرجامعه‌های ۱ تا h باشد:

$$N_1 + N_2 + \dots + N_h = N \quad \text{حجم جامعه}$$

$$n_1 + n_2 + \dots + n_h = n \quad \text{حجم نمونه}$$

انتخاب نمونه از طبقات به روش‌های نمونه‌گیری تصادفی ساده یا سیستماتیک امکان‌پذیر است.

مثال. فرض کنید در بخش جویین از شهرستان سبزوار در استان خراسان رضوی، سه دهستان به نام‌های پیراکوه، بالاجوین و پایین جویین، ۳۰ سکونت‌گاه روستایی وجود دارد و میزان تولید گندم آن‌ها با توجه به موقعیت جغرافیایی، به صورت زیر گزارش شده باشد (ارقام به ده میلیون تن):

دهستان پایین جویین	۵	۶	۴	۷/۵	۸/۵	۶/۷۵	۹	۷	۸	۵/۵
دهستان بالاجوین	۴	۳	۴/۵	۳/۵	۵	۴/۲۵	۳/۷۵	۴/۲۵	۳/۲۵	۶
دهستان پیراکوه	۲/۷۵	۱/۵	۳/۵	۱	۲/۵	۳/۲۵	۳/۷۵	۳	۲	۱/۷۵

فرض کنید بخواهیم از این جامعه آماری یک نمونه ۹ تایی انتخاب کنیم، از آنجا که توزیع تولید در موقعیت‌های جغرافیایی یکسان نیست، بنابراین از هر طبقه یا دهستان، یک نمونه ۳ تایی انتخاب می‌شود:

$$N = 30, h = 3, n_1 = 3, n_2 = 3, n_3 = 3$$

برای انتخاب نمونه از داخل موقعیت‌ها می‌توان از روش نمونه‌گیری تصادفی ساده یا سیستماتیک استفاده کرد.

دهستان پایین جویین	۶	۷/۵	۵/۵
دهستان بالاجوین	۴	۳/۵	۴/۲۵
دهستان پیراکوه	۱/۵	۳/۷۵	۲/۷۵