



دانشگاه علامه طباطبائی
 دانشکده‌ی اقتصاد
 گروه آمار، ریاضی و کامپیوتر
 پایان‌نامه برای دریافت درجه‌ی کارشناسی ارشد آمار اجتماعی و اقتصادی

عنوان

نمونه‌گیری مجموعه‌ی رتبه‌دار میانه‌ای طبقه‌بندی شده

پژوهش‌گر
 ثمینه حاجی قربانی

استاد راهنما
 دکتر روشنگر علی‌اکبری صبا

استاد مشاور
 دکتر نادر نعمت‌الهی
 بهمن ۱۳۹۰

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

کلیه حقوق مادی و معنوی اعم از چاپ و تکثیر، نسخه برداری، ترجمه، اقتباس و ... از این پایان نامه برای دانشگاه علامه طباطبائی محفوظ است. نقل مطلب با ذکر منبع مانعی ندارد.

تأیید پایان‌نامه‌ی کارشناسی ارشد توسط دانشجو

عنوان پایان‌نامه: نمونه‌گیری مجموعه‌ی رتبه‌دار میانه‌ای طبقه‌بندی‌شده

نام دانشجو: ثمینه حاجی‌قربانی

شماره‌ی دانشجویی: ۸۸۱۲۵۱۱۶۱۰۵

استاد راهنما: دکتر روشنگر علی‌اکبری صبا

این جانب ثمینه حاجی‌قربانی دانشجوی کارشناسی ارشد رشته‌ی آمار اجتماعی و اقتصادی دانشکده‌ی اقتصاد دانشگاه علامه طباطبائی گواهی می‌نمایم پژوهش‌های ارائه‌شده در پایان‌نامه با عنوان مذکور توسط شخص این جانب انجام شده است و درستی مطالب نگارش یافته مورد تأیید است. همچنین گواهی می‌نمایم مطالب مندرج در پایان‌نامه تاکنون برای دریافت هیچ نوع مدرک یا امتیازی توسط این جانب یا فرد دیگری در هیچ کجا ارائه نشده است و در نگارش متن پایان‌نامه شیوه‌ی نگارش مصوب دانشکده‌ی اقتصاد را به طور کامل رعایت نموده‌ام. چنانچه در هر زمان خلاف آنچه گواهی نموده‌ام مشاهده گردد خود را از آثار حقیقی و حقوقی ناشی از دریافت مدرک کارشناسی ارشد محروم می‌دانم و هیچ‌گونه ادعایی نخواهم داشت.

امضا دانشجو:

تاریخ:

اندازه

تقدیرم به هر بی آملی که

هتشان بر کرد زم ارت.

سپاس‌گزاری

سپاس خدای را که هر توفیقی در گرو عنایت اوست. اکنون که با یاری او توانسته‌ام تلاشی هر چند ناچیز را در راه کسب دانش به انجام برسانم، بر خود لازم می‌دانم از استاد راهنمای بزرگوارم، سرکار خانم دکتر روشنگر علی‌اکبری صبا، که به پایان رساندن این تحقیق جز با راهنمایی و هدایت‌های بی‌دریغ ایشان میسر نبود، قدردانی نمایم. از استاد مشاورم جناب آقای دکتر نادر نعمت الهی که تذکراتشان باعث غنای پایان‌نامه شد، تشکر می‌نمایم. همچنین از جناب آقای دکتر نواب‌پور و جناب آقای دکتر محمد بامنی مقدم که زحمت داوری این اثر را به عهده داشتند سپاس‌گزارم. در پایان، از خانواده‌ام، به ویژه پدر و مادرم که با حمایت‌های خویش، همواره مرا پشتیبانی کرده‌اند نهایت سپاس و قدرشناسی را دارم.

امیدوارم بتوانم از عهده ادای حق این عزیزان برآیم.

بهمن ۹۰

فهرست مطالب

ب	فهرست مطالب
ث	فهرست جدولها
ج	فهرست شکلها

۱ کلیات

۱.۱	مقدمه	۱
۲.۱	نمونه‌گیری از مجموعه‌ی رتبه‌دار	۵
۳.۱	تاریخچه	۹
۴.۱	هدف‌ها و فصل‌بندی	۱۱

۲ نمونه‌گیری از مجموعه‌ی رتبه‌دار و مجموعه‌ی رتبه‌دار میانه‌ای

۱.۲	مقدمه	۱۴
۲.۲	نمونه‌گیری از مجموعه‌ی رتبه‌دار	۱۵
۱.۲.۲	مقایسه‌ی نمونه‌گیری‌های تصادفی ساده و مجموعه‌ی رتبه‌دار	۱۸
۲.۲.۲	کارایی	۲۰
۳.۲	نمونه‌گیری از مجموعه‌ی رتبه‌دار میانه‌ای	۲۲
۱.۳.۲	کارایی	۲۶
۴.۲	مقایسه روش‌ها	۲۶
۵.۲	خلاصه‌ی فصل	۲۸

۳ نمونه‌گیری از مجموعه‌ی رتبه‌دار طبقه‌بندی شده

۱.۳	مقدمه	۲۹
۲.۳	نمونه‌گیری تصادفی طبقه‌بندی شده	۳۱
۱.۲.۳	برآورد میانگین جامعه‌ای	۳۱
۳.۳	نمونه‌گیری از مجموعه‌ی رتبه‌دار طبقه‌بندی شده	۳۲
۱.۳.۳	برآورد میانگین جامعه‌ای	۳۲
۲.۳.۳	کارایی	۳۴
۳.۳.۳	تخصیص نمونه	۳۴
۱.۳.۳.۳	تخصیص بهین	۳۴
۲.۳.۳.۳	تخصیص متناسب	۳۶
۴.۳	خلاصه‌ی فصل	۳۸

۴ نمونه‌گیری از مجموعه‌ی رتبه‌دار میانه‌ای طبقه‌بندی شده

۱.۴	مقدمه	۳۹
۲.۴	روش انتخاب نمونه	۴۰
۳.۴	برآورد میانگین جامعه‌ای	۴۳
۴.۴	کارایی	۴۳
۵.۴	تخصیص نمونه	۴۷
۱.۵.۴	تخصیص بهین	۴۷
۲.۵.۴	تخصیص متناسب	۴۹
۶.۴	خلاصه‌ی فصل	۵۲

۵ شبیه‌سازی

۱.۵	مقدمه	۵۳
۲.۵	مقایسه کارایی روش‌های نمونه‌گیری طبقه‌بندی شده	۵۴
۳.۵	خلاصه‌ی فصل	۶۷

۶۹ پیش‌نهادها
۷۰	مرجع‌ها
۷۳	واژه‌نامه‌ی فارسی به انگلیسی
۷۷	پیوست الف اثبات برخی از لم‌ها و قضیه‌ها
۸۲	پیوست ب دستور نرم‌افزار SAS در قسمت شبیه‌سازی

فهرست جدول‌ها

۱۵.....	۱.۲ مجموعه‌های مرتب شده در یک چرخه
۱۶.....	۲.۲ نمونه‌ی مجموعه‌ی رتبه‌دار با m چرخه هر یک به اندازه‌ی n
۲۷.....	۳.۲ برآورد کارایی برآورد میانگین جامعه‌ای تحت دو روش RSS و $MRSS$
۴۴.....	۱.۴ برآورد کارایی برآورد میانگین جامعه‌ای با $n=12(n_1=5, n_2=7)$
۴۴.....	۲.۴ برآورد کارایی برآورد میانگین جامعه‌ای با $n=14(n_1=8, n_2=6)$
۴۵.....	۳.۴ برآورد کارایی برآورد میانگین جامعه‌ای با $n=18(n_1=10, n_2=8)$
۴۵.....	۴.۴ برآورد کارایی برآورد میانگین جامعه‌ای با $n=18(n_1=4, n_2=6, n_3=8)$
۴۶.....	۵.۴ برآورد کارایی برآورد میانگین جامعه‌ای با $n=15(n_1=3, n_2=5, n_3=7)$
۵۶.....	۱.۵ برآورد کارایی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع نرمال
۵۷.....	۲.۵ برآورد کارایی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع بتا در حالت متقارن $a = b \geq 1$
۵۹.....	۳.۵ برآورد کارایی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع بتا در حالت $a = 1, b > 1$
۶۰.....	۴.۵ برآورد کارایی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع بتا در حالت $a > 1, b = 1$
۶۱.....	۵.۵ برآورد کارایی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع بتا در حالت $a \neq b$ و $a > 1, b > 1$
۶۲.....	۶.۵ برآورد کارایی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع بتا در حالت $a < 1, b < 1$
۶۳.....	۷.۵ برآورد کارایی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع گاما در حالت $I = 1, a > 1$

- ۸.۵ برآورد اریبی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع گاما
در حالت $I = 1, a > 1$ ۶۵
- ۹.۵ برآورد اریبی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع بتا
در حالت $a > 1, b = 1$ ۶۵
- ۱۰.۵ برآورد اریبی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع بتا
در حالت $a > 1, b > 1$ ۶۶
- ۱۱.۵ برآورد اریبی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع بتا
در حالت $a < 1, b < 1$ ۶۶
- ۱۲.۵ برآورد اریبی برآورد میانگین جامعه‌ای به روش $SMRSS$ نسبت به $SRSS$ و $SSRS$ در توزیع بتا
در حالت $I = 1, a > 1$ ۶۷

فهرست شکل‌ها

۷.....	۱.۱ نمونه‌گیری از مجموعه‌ی رتبه‌دار با اندازه‌ی نمونه‌ای $n=3$ در $m=4$ چرخه
۵۶.....	۵.۱ توزیع نرمال
۵۷.....	۵.۲ توزیع بتا در حالت متقارن $a = b \geq 1$
۵۹.....	۵.۳ توزیع بتا در حالت $a = 1, b > 1$
۶۰.....	۵.۴ توزیع بتا در حالت $a > 1, b = 1$
۶۱.....	۵.۵ توزیع بتا در حالت $a > 1, b > 1$ و $a \neq b$
۶۲.....	۶.۵ توزیع بتا در حالت $a < 1, b < 1$
۶۴.....	۷.۵ توزیع گاما در حالت $I = 1, a > 1$

چکیده

در طراحی آمارگیری‌های نمونه‌ای، استفاده از روش‌های نمونه‌گیری که با صرف هزینه‌ای ناچیز منجر به افزایش کارایی برآوردها می‌شوند از اهمیت زیادی برخوردارند. به کارگیری نمونه‌گیری از مجموعه‌ی رتبه‌دار در برآورد میانگین جامعه‌ای علاوه بر افزایش کارایی، منجر به کاهش هزینه در موقعیت‌هایی می‌شود که واحدهای نمونه‌ای را می‌توان به راحتی رتبه‌بندی کرد. در بسیاری از کاربردهای عملی، رتبه‌بندی کامل متغیر مورد بررسی مشکل است و کاهش کارایی برآورد به دست آمده به مقدار خطا در رتبه‌بندی بستگی دارد. برای به حداقل رساندن این خطا، روش نمونه‌گیری از مجموعه‌ی رتبه‌دار میانه‌ای معرفی شد.

در این پایان‌نامه، ابتدا روش‌های نمونه‌گیری از مجموعه‌ی رتبه‌دار و نمونه‌گیری از مجموعه‌ی رتبه‌دار میانه‌ای ارائه می‌شوند و سپس به مقایسه‌ی کارایی این دو روش نمونه‌گیری در برآورد میانگین جامعه‌ای با نمونه‌گیری تصادفی ساده تحت تعداد مشاهددهای اندازه‌گیری شده‌ی یکسان پرداخته می‌شود. در این پایان‌نامه روش نمونه‌گیری از مجموعه‌ی رتبه‌دار طبقه‌بندی‌شده و نمونه‌گیری از مجموعه‌ی رتبه‌دار میانه‌ای طبقه‌بندی‌شده مورد بررسی قرار می‌گیرند و با به کارگیری تخصیص بهین و متناسب در هر طبقه برای تعیین اندازه‌ی نمونه‌ای و با فرض یکسان بودن هزینه‌ی آمارگیری هر واحد در همه‌ی طبقه‌ها به مقایسه‌ی واریانس برآورد میانگین حاصل از تخصیص بهین و متناسب با مقدارهای متناظر در نمونه‌گیری تصادفی طبقه‌بندی‌شده پرداخته می‌شود و در ادامه واریانس برآورد میانگین حاصل از تخصیص بهین و متناسب با هم مقایسه می‌شوند. با انجام یک شبیه‌سازی تحت برخی از توزیع‌های متقارن و نامتقارن، کارایی نمونه‌گیری از مجموعه‌ی رتبه‌دار میانه‌ای طبقه‌بندی‌شده در برآورد میانگین جامعه‌ای با دو روش نمونه‌گیری از مجموعه‌ی رتبه‌دار طبقه‌بندی‌شده و نمونه‌گیری تصادفی طبقه‌بندی‌شده مقایسه و با تغییر پارامتر توزیع‌ها، تأثیر آن بر کارایی بررسی می‌شود.

واژگان کلیدی. نمونه‌گیری تصادفی ساده، نمونه‌گیری مجموعه‌ی رتبه‌دار، نمونه‌گیری مجموعه‌ی رتبه‌دار میانه‌ای، نمونه‌گیری طبقه‌بندی‌شده، کارایی.

فصل اول

کلیات

۱.۱ مقدمه

از آغاز پیدایش بشر تاکنون، همواره پرسش‌های فراوانی در امور زندگی مطرح بوده است که برای پاسخ به آنها، نیاز به گردآوری اطلاعات است. یکی از روش‌های گردآوری اطلاعات برای انجام تحقیق‌های مورد نیاز سرشماری است. سرشماری به روش آمارگیری گفته می‌شود که در آن همه‌ی واحدهای جامعه‌ای از نظر يك یا چند صفت، مورد مطالعه قرار می‌گیرند. بدیهی است در مواردی که واحدهای جامعه‌ای مورد مطالعه بسیار بزرگ یا نامتناهی است، انجام سرشماری بسیار پرهزینه و زمان‌بر است و حتی در برخی از بررسی‌ها انجام سرشماری موجب از دست رفتن ویژگی صفت مورد مطالعه می‌شود و در نتیجه به کارگیری سرشماری در چنین شرایطی معقول نیست.

به دلیل وجود مشکلات بالا، پژوهش‌گران، گاهی پژوهش خود را محدود به نمونه‌ی کوچکی از جامعه‌ی مورد مطالعه می‌سازند. واضح است که در تعیین نمونه و نحوه‌ی انتخاب آن باید

روش‌هایی را به کار گرفت که نمونه‌ی انتخاب شده بتواند معرف خوبی برای جامعه‌ی خود باشد و با حداقل هزینه‌ی ممکن، دقت مورد نظر را تأمین نماید.

یکی از رایج‌ترین هدف‌های آمارگیری نمونه‌ای، برآورد پارامترهایی مانند میانگین یا نسبت صفت مورد نظر در جامعه است. در این صورت با انتخاب یک نمونه از جامعه و محاسبه‌ی آماره‌ی مربوط، از آن به عنوان برآوردی از پارامتر صفت مورد نظر در جامعه استفاده می‌شود.

در سال‌های گذشته، اطلاعات ما درباره‌ی مشخصه‌های جامعه‌های مورد بررسی اغلب از طریق نمونه‌گیری فراهم شده است، اما تا قبل از سال ۱۹۴۰ میلادی به روش و نظریه نمونه‌گیری توجه چندانی نشده بود. با توسعه‌ی علم احتمال و آمار و ارتباط آن با نمونه‌گیری، شاخه‌ی نمونه‌گیری برای برآورد مشخصه‌های جامعه‌ی تحت مطالعه و بررسی ویژگی‌های این گونه برآوردها به سرعت گسترش یافت به طوری که امروزه در بسیاری از علوم دیگر مانند علوم اجتماعی، اقتصادی، پزشکی، روانشناسی، کشاورزی، آموزشی و ... آمارگیری‌های نمونه‌ای کاربرد وسیعی دارد.

نتیجه‌ی نمونه‌گیری‌ها به شرط رعایت اصول و ضوابط آن، از دقت بالایی برخوردار است. از مزیت‌های استفاده از روش نمونه‌گیری نسبت به سرشماری به مواردی مانند صرفه‌جویی در وقت، انرژی و هزینه و جلوگیری از افزایش احتمال پیش آمدن خطاهای غیرنمونه‌گیری می‌توان اشاره کرد.

به عنوان تأکید بر نقشی که نمونه‌گیری در یک آمارگیری نمونه‌ای ایفا می‌کند، مراحل اصلی یک آمارگیری نمونه‌ای به صورت زیر برشمرده می‌شود:

- تعیین هدف‌های آمارگیری
- تعیین جامعه‌ی آمارگیری
- تهیه چارچوب
- تعیین میزان دقت مطلوب
- انتخاب روش نمونه‌گیری
- تعیین روش‌های اندازه‌گیری یا گردآوری داده‌ها
- تعیین اندازه‌ی نمونه‌ای
- تعیین روش برآورد
- آموزش آمارگیران
- گردآوری داده‌ها
- پردازش داده‌ها
- استخراج اطلاعات و اطلاع رسانی

روش نمونه‌گیری و اطلاعات حاصل از مطالعه‌های پیشین در مورد صفت‌های مورد نظر در جامعه، در تعیین اندازه‌ی نمونه‌ای نقش مهمی دارند و از آنجایی که روش برآورد پارامترها با توجه به روش نمونه‌گیری انتخاب می‌شود، از این رو انتخاب روش نمونه‌گیری مناسب در آمارگیری‌های نمونه‌ای از اهمیت بالایی برخوردار است.

به طور کلی روش‌های نمونه‌گیری را می‌توان به دو بخش نمونه‌گیری احتمالاتی و نمونه‌گیری نااحتمالاتی تقسیم کرد. در نمونه‌گیری احتمالاتی، هر یک از واحدهای جامعه‌ای برای وارد شدن در نمونه از یک احتمال معین (برابر یا نابرابر) و غیرصفر برخوردار است. انتخاب نمونه‌ی احتمالاتی بر اساس عامل شانس یا انتخاب تصادفی انجام می‌گیرد. این عامل شانس یا تصادفی است که تعیین می‌کند کدام واحد باید در نمونه وارد شود. بنابر این، خطاها در نمونه‌گیری احتمالاتی از نوع خطاهای تصادفی قابل اندازه‌گیری هستند.

در نمونه‌گیری نااحتمالاتی، نمونه با قضاوت انسانی انتخاب می‌شود. بنابر این شانس وارد شدن هر یک از واحدهای جامعه‌ای در نمونه، نامعین است. از این رو نمی‌توان خطای تصادفی برآوردگرها را به طور مستقیم محاسبه کرد.

با این‌که استفاده از روش‌های نمونه‌گیری احتمالاتی، پیچیده، زمان‌بر و دارای هزینه‌ای بیش از به‌کارگیری روش‌های نمونه‌گیری نااحتمالاتی است، اما با توجه به این‌که انتخاب واحدها در روش‌های نمونه‌گیری احتمالاتی، تصادفی و دارای احتمال مشخصی است، استفاده از این روش‌ها برآوردهای قابل اعتمادتری از پارامترهای جامعه‌ای ارائه می‌دهد. از مزیت‌های استفاده از نمونه‌گیری احتمالاتی می‌توان به حذف سوگیری آگاهانه یا ناآگاهانه و همچنین امکان محاسبه‌ی مقدار خطا اشاره کرد.

روش‌های رایج مورد استفاده برای نمونه‌گیری احتمالاتی عبارتند از:

- نمونه‌گیری تصادفی ساده
 - نمونه‌گیری خوشه‌ای
 - نمونه‌گیری چند مرحله‌ای
 - نمونه‌گیری طبقه‌بندی شده
 - نمونه‌گیری سیستماتیک
 - نمونه‌گیری با احتمال متناسب با اندازه
- روش‌های رایج برای نمونه‌گیری نااحتمالاتی نیز عبارتند از:
- نمونه‌گیری هدفمند یا با داوری
 - نمونه‌گیری سهمیه‌ای

- نمونه‌گیری داوطلبانه

- نمونه‌گیری گلوله برفی

در ادامه به اختصار، نمونه‌گیری تصادفی ساده به عنوان یکی از پرکاربردترین روش‌های نمونه‌گیری احتمالاتی و نمونه‌گیری هدفمند به عنوان نماینده‌ای از روش‌های نمونه‌گیری نااحتمالاتی شرح داده می‌شوند.

نمونه‌گیری تصادفی ساده

نمونه‌گیری تصادفی ساده روشی کاربردی و پایه‌ای برای نمونه‌گیری است که در آن همه‌ی واحدهای جامعه‌ای، احتمال انتخاب برابر در نمونه دارند. هم‌چنین هر ترکیبی از واحدهای جامعه‌ای با تعداد واحدهای یکسان و مشخص، شانس یکسانی برای انتخاب شدن در نمونه را دارند. در صورتی که پارامتر مورد نظر میانگین باشد، میانگین صفت مورد بررسی در بین واحدهای یک نمونه‌ی تصادفی ساده، برابری ناریب برای میانگین جامعه‌ای است. به عبارت دیگر مقدار مورد انتظار توزیع نمونه‌گیری میانگین در نمونه‌گیری تصادفی ساده با میانگین جامعه‌ای برابر است.

در نمونه‌گیری تصادفی ساده، واریانس برآورد میانگین جامعه‌ای علاوه بر اندازه‌ی نمونه به تغییر پذیری مشخصه‌ی تحت بررسی نیز وابسته است. در صورتی که جامعه خیلی ناهمگن باشد و به دلیل محدودیت‌های اقتصادی نتوان اندازه‌ی نمونه‌ای را بزرگ در نظر گرفت، برآورد میانگین جامعه‌ای که با این روش نمونه‌گیری ارائه می‌شود ممکن است به اندازه‌ی کافی دقیق نباشد. در نتیجه در عمل در صورتی که امکان انتخاب نمونه‌ی بزرگ فراهم نباشد، از نمونه‌گیری تصادفی ساده به تنهایی استفاده نمی‌شود. به عنوان مثال ابتدا با استفاده از اطلاعات کمکی، جامعه طبقه‌بندی و سپس درون هر طبقه از نمونه‌گیری تصادفی ساده استفاده می‌شود. از سوی دیگر گاهی شرایط جامعه و چارچوب ایجاب می‌کند تا از ترکیب روش‌های نمونه‌گیری برای انتخاب نمونه استفاده شود. به عنوان مثال ممکن است از نمونه‌گیری تصادفی ساده برای انتخاب خوشه‌هایی از جامعه استفاده شود یا در نمونه‌گیری‌های چند مرحله‌ای پیچیده از نمونه‌گیری تصادفی ساده تنها در یک یا دو مرحله استفاده شود.

نمونه‌گیری هدفمند یا با داوری

یکی از پرکاربردترین روش‌های نمونه‌گیری نااحتمالاتی که گاهی در عمل به کار می‌روند، نمونه‌گیری هدفمند یا با داوری است. در این روش نمونه‌گیری، واحدهایی انتخاب می‌شوند که به نظر می‌رسند نماینده‌ی مناسبی از واحدهای جامعه‌ای هستند. در نمونه‌گیری هدفمند مطلوبیت واحدهای نمونه‌ای و نماینده بودن آن‌ها به میزان شناخت جامعه و ویژگی‌های آن بستگی دارد. انتخاب نمونه با استفاده از نمونه‌گیری هدفمند به آسانی صورت می‌گیرد. عیب

نمونه‌گیری هدفمند این است که از نظر ریاضی شناختی از قابلیت اعتماد برآوردهای حاصل به دست نمی‌آید.

۲.۱ نمونه‌گیری از مجموعه‌ی رتبه‌دار

نمونه‌گیری تصادفی ساده، رایج‌ترین روش نمونه‌گیری برای گردآوری داده‌ها از جامعه‌ی مورد نظر برای استنباط درباره‌ی پارامتر نامعلومی از آن مانند میانگین است. در حالتی که جامعه خیلی ناهمگن باشد و به دلیل محدودیت‌های اقتصادی نتوان اندازه‌ی نمونه‌ای را بزرگ اختیار کرد ممکن است در نمونه‌های مختلفی که با روش نمونه‌گیری تصادفی ساده انتخاب می‌شود، نوسان‌های شدیدی مشاهده شود. در این صورت با به‌کارگیری روش نمونه‌گیری تصادفی ساده، محقق کنترلی روی انتخاب واحدهای نمونه‌ای ندارد و نمی‌توان برآوردی با دقت کافی، برای پارامتر مورد نظر جامعه ارائه داد. آمارشناسان، روش‌های گوناگونی را برای مقابله با مشاهده‌ی چنین نمونه‌های غیر نماینده‌ای از جامعه ارائه داده‌اند. طرح‌های نمونه‌گیری همچون نمونه‌گیری تصادفی طبقه‌بندی شده، ساختاری روی فرایند نمونه‌گیری برای بهبود احتمال این‌که داده‌های گردآوری شده، نماینده‌ی خوبی از جامعه باشند، ایجاد می‌کنند. در واقع مهم‌ترین هدف در انتخاب روشی برای گردآوری داده‌ها، این است که واحدهای نمونه‌ای نماینده‌ی خوبی برای جامعه باشند و هدف دیگر به حداقل رساندن هزینه‌های مربوط به گردآوری داده‌ها است. این هزینه‌ها، هزینه‌ی مربوط به انتخاب واحدهای نمونه‌ای و اندازه‌گیری صفت واحدهای انتخاب شده از جامعه را شامل می‌شود. در انجام یک آمارگیری، استفاده از روش‌های نمونه‌گیری که با صرف هزینه‌ی ناچیز، منجر به افزایش کارایی برآوردها می‌شوند، از اهمیت زیادی برخوردار است. در آمارگیری‌های نمونه‌ای گاهی می‌توان به آسانی و با صرف هزینه‌ی ناچیز، بدون نیاز به اطلاع دقیق در مورد صفت مورد نظر، واحدهای جامعه‌ای تحت مطالعه را رتبه‌بندی کرد. برای انجام رتبه‌بندی بدون این‌که از مقدار واقعی صفت مورد نظر اطلاعی در دست باشد، اغلب از نظر کارشناسان، مقایسه‌های تصویری و یا مقادیرهای متغیر کمکی که با صفت مورد نظر همبستگی زیادی دارد و اندازه‌گیری آن آسان و کم هزینه است، استفاده می‌شود. در شرایطی که اندازه‌گیری واحدهای جامعه‌ای مشکل، مخرب و یا هزینه‌بر است اما می‌توان واحدهای جامعه‌ای را به راحتی رتبه‌بندی کرد از نمونه‌گیری از مجموعه‌ی رتبه‌دار یا تعمیم‌های آن برای به دست آوردن نمونه‌ی نماینده‌ای از جامعه و برآورد کاراتری از میانگین آن استفاده می‌شود. نمونه‌گیری از مجموعه‌ی رتبه‌دار که توسط مک‌ایتنایر (۱۹۵۲) معرفی شد، روشی است که برای هر دو هدف اشاره شده در انتخاب روش نمونه‌گیری مناسب مورد توجه قرار می‌گیرد، اما در اغلب آمارگیری‌ها هدف از به‌کارگیری نمونه‌گیری از مجموعه‌ی رتبه‌دار به دست آوردن نمونه‌ای است که در مقایسه با روش نمونه‌گیری تصادفی ساده با همان

تعداد نمونه، دامنه‌ی کامل‌تری از مقدارهای جامعه‌ای را در برگیرد. به عبارتی دیگر نمونه‌ی به دست آمده از روش نمونه‌گیری از مجموعه‌ی رتبه‌دار نسبت به روش نمونه‌گیری تصادفی ساده، نماینده‌ی بهتری از جامعه باشد و دامنه‌ی کامل‌تری از واحدهای جامعه‌ای را پوشش دهد که این امر منجر می‌شود با صرف هزینه‌ی مورد نظر، برآورد کاراتری از میانگین جامعه‌ای به دست آید. این افزایش کارایی در نتیجه‌ی رتبه‌بندی واحدها است که با هزینه‌ی ناچیزی صورت می‌گیرد.

در حالت کلی، اگر اندازه‌ی نمونه‌ای n_0 به صورت $n_0 = n \times m$ مورد نظر باشد یعنی انتخاب نمونه‌ها در m چرخه، هر چرخه شامل n واحد صورت گیرد، می‌توان آگوریتم اصلی این روش نمونه‌گیری را به صورت زیر بیان کرد:

مرحله‌ی اول: n^2 واحد نمونه‌ای از جامعه‌ی تحت بررسی به صورت تصادفی انتخاب می‌کنیم.
مرحله‌ی دوم: n^2 واحد انتخابی را به صورت تصادفی به n مجموعه‌ی n تایی اختصاص می‌دهیم.

مرحله‌ی سوم: رتبه‌بندی واحدها در هر یک از مجموعه‌ها را به صورت صعودی بدون داشتن هیچ اطلاعی از مقدار واقعی آن‌ها بر اساس قضاوت شخصی یا با کمک متغیر کمکی انجام می‌دهیم.

مرحله‌ی چهارم: واحد دارای کوچک‌ترین رتبه در اولین مجموعه، واحد دارای دومین رتبه در دومین مجموعه و به همین ترتیب تا واحد دارای بزرگ‌ترین رتبه در آخرین مجموعه را برای تشکیل نمونه بر می‌گزینیم.

مرحله‌ی پنجم: مرحله‌های اول تا چهارم را برای m بار (چرخه) تکرار می‌کنیم تا نمونه‌ای به اندازه‌ی مطلوب $n_0 = n \times m$ به دست آید.

در شکل ۱.۱ نحوه‌ی انتخاب نمونه به صورت تصویری برای اندازه‌ی نمونه‌ای $n=3$ در $m=4$ چرخه نشان داده شده است.

چرخه	رتبه		
	1	2	3
1	1	0	0
	0	1	0
	0	0	1
2	1	0	0
	0	1	0
	0	0	1
3	1	0	0
	0	1	0
	0	0	1
4	1	0	0
	0	1	0
	0	0	1

شکل ۱.۱: نمونه‌گیری از مجموعه‌ی رتبه‌دار با اندازه‌ی نمونه‌ای $n=3$ در $m=4$ چرخه

در ادامه با بررسی چند مثال با کاربردهای نمونه‌گیری از مجموعه‌ی رتبه‌دار آشنا می‌شویم. فرض کنید می‌خواهیم متوسط قد دانشجویان یک کلاس را با یک نمونه‌ی تصادفی با اندازه‌ی نمونه‌ای سه دانشجو، برآورد کنیم. ساده‌ترین راه برای به‌دست آوردن نمونه، انتخاب سه دانشجو به تصادف از جامعه‌ی دانشجویان کلاس مورد نظر و سپس اندازه‌گیری قد آن‌ها است. اما از آنجایی که در روش نمونه‌گیری تصادفی ساده ممکن است بدنمونه روی دهد یا به عبارت دیگر کنترلی روی واحدهای انتخاب شده وجود ندارد و ممکن است از سه دانشجوی انتخاب شده، دو نفر خیلی کوتاه قد و یک نفر خیلی بلند قد باشند و یا ممکن است هر سه نفر خیلی بلند قد باشند.