

# روش های نمونه گیری مبتنی بر مجموعه های رتبه دار و کاربرد آن ها در نمونه گیری طبقه بندی شده

محرر طرح: محمد صفری جوزانی

محمد حضرتی

امیر حسین میرزایی

فرشید جمشیدی

گروه پژوهشی طرح های فنی و روش های آماری

پژوهشکده های آمار

## بهنام خداوند جان و خرد

### پیش گفتار

پژوهشکده‌ی آمار، به منظور دستیابی به مناسب‌ترین فنون و روش‌های تولید آمار، هر ساله تعدادی طرح‌های پژوهشی بویژه در ارتباط با فعالیت‌های مرکز آمار ایران با در نظر گرفتن اولویت موضوع، در دستور کار خود قرار می‌دهد. از جمله موضوعاتی که اخیراً با تاکید بیش‌تری مطرح است به‌کارگیری روش‌های نمونه‌گیری جدید و مناسب برای افزایش کارایی و دقت برآوردهای بدست آمده است.

روش‌های نمونه‌گیری مجموعه رتبه‌دار که اغلب منجر به برآوردهای با دقت بیش‌تری نسبت به برآوردهای متناظر تحت نمونه‌گیری تصادفی ساده با همان مقدار عضو اندازه‌گیری شده می‌باشد، یکی از روش‌های نمونه‌گیری است که در این طرح پژوهشی به آن پرداخته می‌شود که به کمک این روش می‌توان استنباط‌های دقیق‌تری درباره جمعیت مورد مطالعه به دست آورد.

در گروه مطالعاتی طرح مذکور، آقای دکتر محمد جعفری جوزانی به عنوان مجری طرح و آقایان محمد حضرتی، امیر حسین میرزایی و فرشید جمشیدی به عنوان همکار عضویت داشتند که بدین‌وسیله از زحمات یکایک این افراد تشکر و قدردانی می‌شود.

در طول انجام این طرح پژوهشی سرکار خانم مهنوش میرمحمد با دقت و حوصله بسیار زیاد زحمت حروف چینی و تایپ مستندات طرح را به عهده داشته‌اند که بدین‌وسیله از تلاش‌های بی‌وقفه ایشان تشکر و قدردانی می‌شود.

گروه پژوهشی طرح‌های فنی و روش‌های آماری

## پیش‌گفتار مجری

گاهی اوقات واحدهای برخی از جوامع تحت بررسی را می‌توان به آسانی و با صرف کم‌ترین هزینه بدون آن‌که اطلاع دقیقی از مقدار خصوصیت مورد بررسی در دست باشد رتبه‌بندی کرد که برای این کار اغلب از نظرات کارشناسی یا مقادیر متغیرهای کمکی مرتبط با خصوصیت مورد بررسی که اندازه‌گیری آن‌ها آسان و کم هزینه باشد استفاده می‌شود. این وضعیت به ویژه در بررسی‌های نمونه‌ای زیست‌محیطی و آزمایشگاهی، مسائل مربوط به جنگل‌داری، برآورد میزان مصرف یک مقوله مورد نظر یا میزان محصول و ... اتفاق می‌افتد. در واقع در شرایطی که اندازه‌گیری واحدهای جامعه مشکل یا پرهزینه باشد اما بتوان واحدهای جامعه را به راحتی رتبه‌بندی کرد نمونه‌گیری مجموعه رتبه‌دار و تعمیم‌های آن روش‌های کاراتری را برای برآورد میانگین و گاهی اوقات واریانس جامعه مورد مطالعه ارائه می‌کنند. در روش‌های نمونه‌گیری مجموعه رتبه‌دار اغلب برآوردهای حاصل دارای دقت بیشتری نسبت به برآوردهای متناظر تحت نمونه‌گیری تصادفی ساده با همان مقدار عضو اندازه‌گیری شده می‌باشند و لذا به کمک آن‌ها می‌توان استنباط‌های دقیق‌تری درباره جمعیت مورد مطالعه انجام داد.

در این پژوهش ابتدا در فصل اول سه روش نمونه‌گیری مجموعه رتبه‌دار، مجموعه رتبه‌دار با انتخاب‌های تصادفی و مجموعه رتبه‌دار کامل را در برآورد میانگین جامعه مورد مطالعه قرار می‌دهیم. سپس در فصل دوم با تعمیم نتایج فصل اول چگونگی استفاده از این روش‌ها را در برآورد (گشتاوری) بعضی از پارامترهای جمعیت و با رویکرد ناپارامتری مورد مطالعه قرار می‌دهیم. همچنین کاربرد جدیدی از نتایج به‌دست آمده در برآورد هسته‌ای تابع چگالی احتمال را بیان کرده و با انجام یک مطالعه شبیه‌سازی شده مقایسه‌ای بین برآوردهای به‌دست آمده از سه روش مذکور در برآورد هسته‌ای تابع چگالی احتمال بعضی از توزیع‌های مهم آماری انجام می‌دهیم. در فصل

سوم با رویکرد پارامتری و در برآورد میانگین توزیع نمایی و خانواده توزیع‌های تابع توان با ارائه برآوردگرهای نارایب میانگین به کمک روش‌های نمونه‌گیری مجموعه رتبه‌دار تصادفی، مجموعه رتبه‌دار کامل و با معرفی بعضی روش‌های جدید مقایسه‌ای بین این برآوردگرها و برآوردگرهای موجود در مطالعات مشابه انجام می‌دهیم. در هر مورد نشان می‌دهیم در برآورد میانگین جامعه استفاده هم‌زمان توزیع صفت مورد بررسی و روش‌های نمونه‌گیری مبتنی بر مجموعه رتبه‌دار برآوردگرهای کاراتری در مقایسه با برآوردگرهای به‌دست آمده به کمک روش نمونه‌گیری تصادفی ساده، که در عمل از آن استفاده می‌شود، ارائه می‌دهد. نتایج تحقیقات مشابه نشان می‌دهد استفاده از روش‌های نمونه‌گیری مجموعه رتبه‌دار و خواص آن‌ها در طرح‌های نمونه‌گیری چند مرحله‌ای که به ویژه در مراحل نهایی آن‌ها و در انتخاب اعضای نمونه از روش تصادفی ساده استفاده می‌شود (مانند نمونه‌گیری طبقه‌بندی شده، نمونه‌گیری گروهی و ...) دقت روش نمونه‌گیری چند مرحله‌ای به دست آمده را در برآورد خصیصه مورد نظر افزایش می‌دهد. برای این منظور در فصل چهارم روش نمونه‌گیری طبقه‌بندی مجموعه رتبه‌دار معمولی را توضیح داده و دو روش جدید نمونه‌گیری طبقه‌بندی مجموعه رتبه‌دار تصادفی و نمونه‌گیری طبقه‌بندی مجموعه رتبه‌دار کامل را معرفی می‌کنیم. سپس خواص آن‌ها را در برآورد بعضی از پارامترهای جمعیت مورد مطالعه قرار می‌دهیم. در نهایت کاربردی از نتایج به دست آمده را در برآورد میزان مصرف بنزین ارائه می‌کنیم.

با تشکر

محمد جعفری جوزانی

## فهرست مطالب

صفحه	عنوان
۱	فصل اول: روش های نمونه گیری مبتنی بر مجموعه های رتبه دار
۱	۱-۱) مقدمه .....
۵	۲-۱) تاریخچه .....
۱۱	۳-۱) مفاهیم اولیه .....
۱۲	۱-۳-۱) مرتب سازی غیرکامل نسبت به متغیر مورد بررسی .....
۱۳	۲-۳-۱) مرتب سازی به کمک متغیر کمکی .....
۱۴	۴-۱) روش نمونه گیری مجموعه رتبه دار ساده .....
۱۶	۵-۱) روش نمونه گیری مجموعه رتبه دار تصادفی .....
۱۸	۶-۱) روش نمونه گیری مجموعه رتبه دار کامل .....
۲۳	فصل دوم: برآورد میانگین ها به کمک روش های نمونه گیری مبتنی بر مجموعه های رتبه دار
۲۴	۱-۲) برآورد میانگین ها به کمک نمونه گیری مجموعه رتبه دار .....
۲۹	۲-۲) برآورد میانگین ها به کمک نمونه گیری مجموعه رتبه دار تصادفی .....
۳۵	۳-۲) برآورد میانگین به کمک روش نمونه گیری مجموعه رتبه دار کامل .....
۴۰	۴-۲) برآورد هسته ای تابع چگالی به کمک روش های نمونه گیری مجموعه رتبه دار .....
۴۲	۱-۴-۲) برآورد هسته ای تابع چگالی به کمک روش های نمونه گیری مجموعه رتبه دار .....
۴۵	۲-۴-۲) برآورد هسته ای تابع چگالی به کمک روش نمونه گیری مجموعه رتبه دار تصادفی .....
۵۱	۳-۴-۲) برآورد هسته ای تابع چگالی به کمک روش نمونه گیری مجموعه رتبه دار کامل .....
۵۲	۴-۴-۲) مقایسه برآوردهای هسته ای تابع چگالی از طریق شبیه سازی .....
۵۵	۱-۴-۴-۲) توزیع نرمال استاندارد .....
۶۲	۲-۴-۴-۲) توزیع نرمال آمیخته .....
۶۷	۳-۴-۴-۲) توزیع لگ نرمال .....

فصل سوم: برآوردیابی پارامتری به کمک روش‌های نمونه‌گیری مبتنی بر مجموعه رتبه‌دار ۷۷

- ۷۸ ..... (۱-۳) مقدمه
- ۸۰ ..... (۲-۳) برآوردیابی میانگین توزیع نمایی
- ۸۲ ..... (۱-۲-۳) روش نمونه‌گیری مجموعه رتبه‌دار
- ۸۴ ..... (۲-۲-۳) روش نمونه‌گیری مجموعه رتبه‌دار با انتخاب‌های تصادفی
- ۸۷ ..... (۳-۲-۳) روش نمونه‌گیری مجموعه رتبه‌دار میانه‌ای
- ۸۹ ..... (۴-۲-۳) روش نمونه‌گیری مجموعه رتبه‌دار فرین
- ۹۲ ..... (۵-۲-۳) روش نمونه‌گیری مجموعه رتبه‌دار کامل و تعمیم‌های آن
- ۹۵ ..... (۶-۲-۳) مقایسه
- ۹۸ ..... (۳-۳) برآورد میانگین در خانواده توزیع‌های توانی
- ۱۰۰ ..... (۱-۳-۳) روش نمونه‌گیری مجموعه رتبه‌دار معمولی
- ۱۰۱ ..... (۲-۳-۳) روش نمونه‌گیری مجموعه رتبه‌دار تصادفی
- ۱۰۲ ..... (۳-۳-۳) روش نمونه‌گیری مجموعه رتبه‌دار میانه‌ای
- ۱۰۳ ..... (۴-۳-۳) روش نمونه‌گیری مجموعه رتبه‌دار فرین
- ۱۰۴ ..... (۵-۳-۳) روش نمونه‌گیری مجموعه رتبه‌دار کامل و تعمیم آن
- ۱۰۶ ..... (۶-۳-۳) مقایسه

فصل چهارم: روش‌های نمونه‌گیری طبقه‌بندی مجموعه رتبه‌دار ۱۱۳

- ۱۱۴ ..... (۱-۴) مقدمه
- ۱۱۶ ..... (۲-۴) نمونه‌گیری طبقه‌بندی مجموعه رتبه‌دار
- ۱۱۶ ..... (۱-۲-۴) نمادگذاری
- ۱۱۷ ..... (۲-۲-۴) برآورد میانگین
- ۱۲۱ ..... (۳-۴) نمونه‌گیری طبقه‌بندی مجموعه رتبه‌دار تصادفی
- ۱۲۲ ..... (۱-۳-۴) نمادگذاری
- ۱۲۳ ..... (۲-۳-۴) برآورد میانگین
- ۱۲۶ ..... (۳-۴) نمونه‌گیری طبقه‌بندی مجموعه رتبه‌دار کامل

۱۲۷	..... (۱-۳-۴) نمادگذاری
۱۲۸	..... (۲-۳-۴) برآورد میانگین
۱۳۳	..... <b>فصل پنجم: برآورد میزان مصرف بنزین کشور</b>
۱۳۴	..... (۱-۵) ضرورت تحقیق
۱۳۶	..... (۲-۵) مقدمات
۱۳۶	..... (۱-۲-۵) جامعه مورد نمونه‌گیری
۱۴۰	..... (۲-۲-۵) اندازه نمونه
۱۴۱	..... (۳-۵) برآورد به روش نمونه‌گیری تصادفی ساده
۱۴۲	..... (۴-۵) برآورد به روش نمونه‌گیری مجموعه رتبه‌دار معمولی
۱۴۴	..... (۵-۵) برآورد به روش نمونه‌گیری مجموعه رتبه‌دار تصادفی
۱۴۶	..... (۶-۵) برآورد به روش نمونه‌گیری مجموعه رتبه‌دار کامل
۱۴۷	..... (۷-۵) نتیجه‌گیری و مقایسه
۱۵۱	..... <b>ضمیمه</b>
۱۵۹	..... <b>مراجع</b>

## فصل اول

روش‌های نمونه‌گیری مبتنی بر مجموعه‌های رتبه‌دار



# فصل اول

## روش‌های نمونه‌گیری مبتنی بر مجموعه‌های رتبه‌دار

در این فصل ابتدا مقدمه‌ای درباره روش‌های نمونه‌گیری مجموعه رتبه‌دار و تاریخچه مختصری از آن شامل اهداف، مطالعات انجام شده و کاربردهای آن بیان می‌کنیم. سپس ضمن بیان مقدمات لازم، سه روش نمونه‌گیری مجموعه رتبه‌دار معمولی، مجموعه رتبه‌دار با انتخاب‌های تصادفی و مجموعه رتبه‌دار کامل را بیان کرده، روش انجام آن‌ها و خواص برآوردهای به‌دست آمده توسط این سه روش برای میانگین جامعه را مورد مطالعه قرار می‌دهیم.

## ۱-۱) مقدمه

هرگاه بخواهیم درباره یک جمعیت یا خصوصیتی مورد نظر از آن جمعیت به طریق عملی رفتار کرده و از اعداد به دست آمده یاری بگیریم بایستی گام‌های زیر طی شوند:

۱) انتخاب روش مناسب برای کسب اطلاعات از خصوصیات مورد مطالعه

۲) کسب اطلاعات لازم وابسته به خصوصیت مورد مطالعه از راه بررسی یا طرح آزمایش

۳) اندازه‌گیری اطلاعات جمع‌آوری شده به وسیله مقیاس‌های مناسب و بیان آن‌ها به صورت

عددی به نام داده‌ها

۴) تنظیم و تحلیل داده‌ها طبق اصول و ضوابطی معقول

برای دستیابی به نمونه‌ای که نماینده جمعیت مورد مطالعه باشد روش‌های مختلف نمونه‌گیری وجود دارد. برای مثال یکی از ساده‌ترین این روش‌ها، استفاده از روش نمونه‌گیری تصادفی ساده است که اغلب از آن برای برآورد میانگین مشخصه‌های مورد مطالعه جمعیت در مواقعی استفاده می‌شود که از جمعیت مورد نظر هیچ‌گونه شناخت و اطلاعی در دست نیست. دقت به‌کارگیری این روش در برآورد میانگین جامعه علاوه بر آن که به اندازه نمونه بستگی دارد، به همگنی یا ناهمگنی جامعه از لحاظ خصوصیت مورد بررسی نیز وابسته است. در این روش در اکثر موارد اگر نتوان اندازه‌ی نمونه را به اندازه‌ی کافی بزرگ اختیار کرد، آن‌گاه نمی‌توان برآوردی به اندازه‌ی کافی دقیق و مطلوب پیدا کرد. در انجام یک نمونه‌گیری هدف اصلی ارائه روش‌های نمونه‌گیری، نمونه‌ها و برآوردگرهایی است

که به کمک آن بتوان با صحت و دقت بیشتر تخصیص‌های مورد نظر در جامعه را با در نظر گرفتن عواملی مانند هزینه، بودجه، زمان دستیابی به برآوردها و ... برآورد کرد.

دو روش اساسی برای نیل به این مقصود عبارتند از استفاده از برآوردهای پیچیده‌تر یا استفاده از روش‌های نمونه‌گیری مناسب‌تر و یا گاهی پیچیده‌تر. بنابراین ارائه روش‌های جدید نمونه‌گیری برای افزایش کارایی و دقت برآوردهای به دست آمده یکی از مهم‌ترین اهداف نظریه نمونه‌گیری است.

گاهی اوقات واحدهای برخی از جوامع تحت مطالعه را می‌توان به آسانی و با صرف کم‌ترین هزینه، بدون آن که اطلاع دقیقی از مقدار خصوصیت مورد بررسی در دست باشد، رتبه‌بندی کرد. برای این کار اغلب از نظر کارشناسان یا مقادیر متغیری مرتبط با خصوصیت مورد بررسی که اندازه‌گیری آن آسان و کم هزینه باشد استفاده می‌شود. این وضعیت به ویژه در بررسی‌های نمونه‌ای زیست محیطی، آزمایشگاهی، مسائل مربوط به جنگل‌داری، برآورد میزان مصرف یک مقوله مورد نظر یا میزان محصول و ... اتفاق می‌افتد. در واقع در شرایطی که اندازه‌گیری واحدهای جامعه مشکل یا پر هزینه باشد اما بتوان واحدهای جامعه را به راحتی رتبه‌بندی کرد، روش نمونه‌گیری مجموعه رتبه‌دار و تعمیم‌های آن، روش‌های کاراتری را برای برآورد میانگین جامعه مورد مطالعه ارائه می‌کنند. کاربرد عملی روش‌های نمونه‌گیری مبتنی بر مجموعه رتبه‌دار در مواقعی است که بخواهیم بدون صرف هزینه اضافی برآوردهای دقیق‌تری از میانگین (یا سایر پارامترهای جامعه) به دست آوریم. در روش‌های نمونه‌گیری مجموعه رتبه‌دار برآوردها حاصل (برای میانگین جامعه) دارای دقت بیشتری

نسبت به برآوردگر حاصل از نمونه‌گیری تصادفی ساده با همان مقدار عضو اندازه‌گیری شده می‌باشد، بنابراین می‌توان استنباط‌های مناسب‌تری درباره جمعیت مورد مطالعه انجام داد. این افزایش کارایی در نتیجه رتبه‌بندی واحدها، که با کم‌ترین هزینه صورت گرفته، حاصل می‌شود. لازم به ذکر است این روش علاوه بر افزایش دقت ممکن است منجر به صرف هزینه بسیار زیاد شود. برای مثال در بررسی‌های ماهیانه ای که از زمین‌های اطراف یک نیروگاه هسته‌ای به عمل می‌آید تا مشخص شود که آلاینده‌های خاک اطراف نیروگاه به بعضی مواد رادیواکتیویته چگونه است و آیا از نظر مقدار آلاینده‌ها در سطح نرمالی به سر می‌برد، ابتدا نمونه‌هایی انتخاب شده و سپس با انجام آزمایش‌های بسیار پرهزینه مقدار متغیرهای مورد توجه را در آن اندازه‌گیری می‌کنند. لذا مسئله مهم آن است که نمونه انتخاب شده نماینده جمعیت مورد مطالعه باشد. برای این منظور یک روش کارآمد می‌تواند استفاده از نمونه‌گیری مجموعه رتبه‌دار باشد. در واقع دستگاه‌هایی وجود دارند که به طور تقریبی و تنها با حرکت دادن آن‌ها روی سطح خاک می‌توانند تقریبی از مقدار متغیر مورد بررسی در خاک یا قطعه زمین مورد نظر ارائه دهند. بنابراین در این حالت با انتخاب نمونه‌هایی که آغشته بودن آن‌ها به اثر تشعشعات مورد نظر بیشتر است گزینش شده و سپس برای اندازه‌گیری نهایی به آزمایشگاه ارسال می‌شوند. در این طرح پژوهشی ضمن بیان روش‌های نمونه‌گیری مجموعه رتبه‌دار، مجموعه رتبه‌دار با انتخاب‌های تصادفی و رتبه‌دار کامل کاربردهایی از این روش‌ها در برآورد میانگین‌ها و به ویژه میانگین جامعه و همچنین برآورد هسته‌ای تابع چگالی ارائه داده مقایسه‌ای بین آن‌ها انجام می‌دهیم. همچنین برآورد پارامترهای نامعلوم در خانواده‌ای از توزیع‌ها را به کمک روش نمونه‌گیری

مجموعه رتبه‌دار و روش نمونه‌گیری مجموعه رتبه‌دار کامل به‌دست آورده و مقایسه‌ای بین آن‌ها و برآوردهای رایج آن پارامترها ارائه می‌دهیم. با توجه به قراین موجود به نظر می‌رسد استفاده از روش‌های نمونه‌گیری مجموعه رتبه‌دار و خواص آن‌ها در طرح‌های نمونه‌گیری چند مرحله‌ای که به ویژه در مراحل نهایی آن‌ها در انتخاب اعضای نمونه از روش‌های تصادفی ساده استفاده می‌کنند (مانند روش نمونه‌گیری طبقه‌بندی شده با انتخاب‌های تصادفی در طبقات یا روش نمونه‌گیری خوشه‌ای) دقت روش نمونه‌گیری چند مرحله‌ای به‌دست آمده را در برآورد پارامتر مورد نظر افزایش دهد. برای این منظور و برای بررسی این فرضیه سه روش نمونه‌گیری شامل روش نمونه‌گیری طبقه‌بندی مجموعه رتبه‌دار، روش نمونه‌گیری طبقه‌بندی مجموعه رتبه‌دار تصادفی (با انتخاب تصادفی) و روش نمونه‌گیری طبقه‌بندی مجموعه رتبه‌دار کامل را معرفی کرده و دقت برآوردهای به دست آمده برای میانگین جامعه از سه روش فوق را با روش نمونه‌گیری طبقه‌بندی مقایسه می‌کنیم. لازم به ذکر است با توجه به کاربردهای متصور برای این روش به نظر می‌رسد در مرکز آمار ایران و به ویژه در دفتر مطالعات کشاورزی و دفتر مطالعات اقتصاد، صنعت و محیط زیست بتوان از این روش نمونه‌گیری به طور بهینه و شایسته‌ای استفاده کرد.

## ۲-۱) تاریخچه

برای اولین بار روش نمونه‌گیری مجموعه رتبه‌دار توسط مک اینتایر<sup>۱</sup> در سال ۱۹۵۲ به عنوان روشی برای افزایش دقت برآورد میانگین محصول مراتع بدون آن که توجهی به اریبی گزینش نمونه‌های

---

<sup>۱</sup> McIntyre

نماینده جمعیت صورت گرفته باشد ارائه شد. این روش برای افزایش دقت برآورد میانگین و به عنوان رقیبی برای روش نمونه‌گیری تصادفی ساده به‌کار گرفته شد. وی روش نمونه‌گیری مجموعه رتبه‌دار را بر اساس رتبه‌بندی کامل یعنی زمانی که هیچ خطایی در رتبه‌بندی واحدهای نمونه‌ای رخ نمی‌دهد ارائه کرد و برآورد میانگین جامعه را بر اساس این روش با روش نمونه‌گیری تصادفی ساده مورد مقایسه قرار داد. وی همچنین توضیح می‌دهد چگونه ممکن است خطای مرتب‌سازی بر کارایی نمونه‌گیری مجموعه رتبه‌دار تأثیر گذار باشد. برای این منظور وی پیشنهاد می‌دهد که در عمل از مقدار نمونه‌های کوچک‌تری در هر مرحله نمونه‌گیری استفاده شود به گونه‌ای که بتوان به راحتی واحدهای نمونه‌ای را مرتب کرد.

هالز<sup>۱</sup> و دل<sup>۲</sup> در سال ۱۹۶۶ با استفاده از این روش میانگین محصول علوفه و میانگین ارتفاع درختان یک جنگل را برآورد کردند. تاکاهاشی<sup>۳</sup> و واکیموتو<sup>۴</sup> (۱۹۶۸) اولین افرادی بودند که توانستند تئوری روش نمونه‌گیری مجموعه رتبه‌دار را به طور کامل و مبسوط بیان کرده و از آن در اثبات کارایی برآورد میانگین به روش نمونه‌گیری مجموعه رتبه‌دار در مقایسه با روش تصادفی ساده و بعضی از مواردی که در مقاله مک ایتتایر (۱۹۵۲) به صورت فرضیه و بدون ارائه اثبات بیان شده بود استفاده کنند.

آنها میانگین حسابی مشاهدات حاصل از روش نمونه‌گیری مجموعه رتبه‌دار را به عنوان برآوردگر نارایی برای میانگین جامعه ارائه کردند و نشان دادند که این برآوردگر از برآوردگر میانگین جامعه در حالت نمونه‌گیری تصادفی ساده با اندازه نمونه برابر دارای واریانس کم‌تری است. همچنین کران بالایی

---

<sup>1</sup> Halls

<sup>2</sup> Dell

<sup>3</sup> Takahasi

<sup>4</sup> Wakimoto

را برای کارایی یا در واقع عملکرد  $\hat{\mu}_{RSS}$  نسبت به  $\hat{\mu}_{SRS}$  ارائه کردند که در آن منظور از RSS<sup>۱</sup> و SRS<sup>۲</sup> نمونه‌گیری مجموعه رتبه‌دار و نمونه‌گیری تصادفی ساده می‌باشد و نشان دادند:

$$1 \leq RP = \frac{Var(\hat{\mu}_{SRS})}{Var(\hat{\mu}_{RSS})} \leq \frac{k+1}{2}$$

که در آن  $k$  اندازه نمونه استخراج شده در هر مرحله است. لازم به ذکر است این کران بالا اولین بار توسط مک ایتنایر (۱۹۵۲) و در مثال‌های عددی ارائه شده توسط ایشان مشاهده شد ولی در آن زمان وی موفق به اثبات آن در حالت کلی نشد. پس از ارائه پایه تئوری روش نمونه‌گیری مجموعه رتبه‌دار، استفاده از آن به سرعت رونق گرفت و از آن در بسیاری از مطالعات استفاده شد. همان طور که می‌توان حدس زد رتبه‌بندی کامل متغیر مورد بررسی در بسیاری از کاربردهای عملی مشکل یا تقریباً غیر ممکن است. دل و کلاتر<sup>۳</sup> در سال ۱۹۷۲ این روش نمونه‌گیری را وقتی که خطای رتبه‌بندی وجود دارد، مورد استفاده قرار داده و نشان دادند صرف نظر از آن که در رتبه‌بندی واحدهای نمونه‌ای مرتکب خطا شده باشیم یا نه، میانگین نمونه مجموعه رتبه‌دار یک برآورد ناریب برای میانگین جامعه است. دیوید<sup>۴</sup> و لوین<sup>۵</sup> در سال ۱۹۷۲ در یک مطالعه هم‌زمان با دل و کلاتر (۱۹۷۲) یک روش تحلیلی برای محاسبه دقت نسبی روش مجموعه رتبه‌دار در مقایسه با روش تصادفی ساده ارائه کردند و به صورت تئوری نتایج شبیه‌سازی‌های ارائه شده توسط دل و کلاتر را اثبات نمودند. در برخی از مسائل کاربردی رتبه‌بندی متغیر مورد بررسی بدون اندازه‌گیری آن امکان پذیر نیست. در چنین حالتی ممکن است یک متغیر کمکی وجود داشته باشد که با متغیر مورد نظر دارای همبستگی قوی‌ای بوده و ضمناً به آسانی قابل

<sup>۱</sup> Ranked set Sampling (RSS)

<sup>۲</sup> Simple Random Sampling (SRS)

<sup>۳</sup> Clutter

<sup>۴</sup> David

<sup>۵</sup> Levine

اندازه‌گیری است. در این صورت برای رتبه‌بندی واحدهای نمونه‌ای می‌توان آن‌ها را بر اساس متغیر کمکی یا همراه رتبه‌بندی کرد.

اکانل<sup>۱</sup> (۱۹۷۴) در رساله دکتری خود و استوکس<sup>۲</sup> در سال ۱۹۷۷ طی مقاله‌ای کاربرد متغیرهای کمکی در روش نمونه‌گیری مجموعه رتبه‌دار را مطرح کردند. آن‌ها نشان دادند برآوردگر میانگین جمعیت حاصل از روش نمونه‌گیری مجموعه رتبه‌دار کمکی بسیار کارتر از برآورد میانگین نمونه‌گیری تصادفی ساده است. استوکس در سال ۱۹۷۶ و در رساله دکتری خود تحت عنوان «نگاهی به نتایج نمونه‌گیری مجموعه رتبه‌دار» برآوردیابی پارامتری و ناپارامتری پارامترهای مورد نظر را به کمک روش نمونه‌گیری مجموعه رتبه‌دار متعادل و نامتعادل مورد بررسی قرار داد. همچنین در این رساله آزمون ناپارامتری من-ویتنی مجموعه رتبه‌دار معرفی گردید و کارایی آن با آزمون ناپارامتری من-ویتنی معمولی مورد مقایسه قرار گرفت. وی نشان داد که فرم مجموعه رتبه‌دار آزمون من-ویتنی دارای کارایی بسیار بیشتری در مقایسه با فرم معمولی آن است. در سال ۱۹۸۰، استوکس بدون توجه به حضور خطای مرتب‌سازی برآوردگری برای واریانس جمعیت  $\sigma^2$  معرفی کرد که به طور جانبی نااریب بوده و ضمناً کارایی مجانبی آن در مقایسه با برآوردگرهای معمولی  $\sigma^2$  با اندازه برابر، بیش‌تر است.

وی همچنین مطالعاتی در زمینه برآورد ضریب همبستگی در جوامع نرمال دو متغیره به کمک نمونه‌های مجموعه رتبه‌دار انجام داد. در سال ۱۹۹۳، پتیل<sup>۳</sup>، سینها<sup>۴</sup> و تایللی<sup>۵</sup> نشان دادند که متغیر کمکی را می‌توان هم در مرحله‌ی نمونه‌گیری و هم در مرحله‌ی برآوردیابی به طور مؤثر مورد استفاده قرار داد.

---

<sup>1</sup> O'connell

<sup>2</sup> Stokes

<sup>3</sup> Patil

<sup>4</sup> Sinha

<sup>5</sup> Taili



به همین منظور آن‌ها برآوردگر رگرسیون نمونه‌گیری مجموعه رتبه‌دار را با برآوردگر رگرسیونی نمونه‌گیری تصادفی ساده مورد مقایسه قرار دادند. همچنین یو<sup>۱</sup> و لام<sup>۲</sup> در سال ۱۹۹۷ برآوردگر رگرسیونی بر اساس نمونه‌گیری مجموعه رتبه‌دار را در نظر گرفته و آن را با برآوردگرهای عادی و رگرسیونی تصادفی ساده مقایسه کردند. در سال ۱۹۹۶، سینها و همکارانش برآورد میانگین و واریانس توزیع نرمال و میانگین توزیع نمایی را مورد مطالعه قرار داده و با انجام اصلاحاتی روی برآوردهای حاصل آن‌ها را بهبود بخشیدند. در سال ۱۹۹۹، یو، لام و سینها روش‌هایی را برای برآورد واریانس توزیع نرمال در نمونه‌گیری مجموعه رتبه‌دار در دو حالت متعادل و نامتعادل مورد مطالعه قرار داده و کارایی برآوردهای حاصل را با هم مقایسه کردند. تیواری<sup>۳</sup> و کوام<sup>۴</sup> در سال ۲۰۰۱ برآوردهایی برای میانگین و واریانس جامعه در روش نمونه‌گیری مجموعه رتبه‌دار برای خانواده توزیع‌های مکان-مقیاس به‌دست آورده و کارایی آن‌ها را با هم مقایسه نمودند. همچنین مک ایچرن<sup>۵</sup> و همکاران در سال ۲۰۰۲ برآوردگر نارایب جدیدی برای واریانس جمعیت بر اساس نمونه‌گیری مجموعه رتبه‌دار ارائه کردند. آن‌ها نشان دادند که این برآوردگر جدید بسیار کاراتر از برآوردگر معمولی واریانس و برآوردگر پیشنهادی توسط استوکس است. برای مطالعه نسبتاً کامل عناوین و خلاصه تمامی مقالات چاپ شده در رابطه با روش‌های نمونه‌گیری مجموعه رتبه‌دار تا سال ۱۹۹۵ به مقاله کاور<sup>۶</sup> و همکاران مراجعه کنید. همچنین برای مطالعه بیشتر در زمینه برآوردیابی پارامتری و ناپارامتری به

---

<sup>1</sup> Yu

<sup>2</sup> Lam

<sup>3</sup> Tiwari

<sup>4</sup> Kvam

<sup>5</sup> MacEachern

<sup>6</sup> Kaur

کمک روش نمونه‌گیری مجموعه رتبه‌دار و همچنین تحلیل رگرسیونی و آزمون‌های ناپارامتری به کمک این روش به چن<sup>۱</sup> و همکاران (۲۰۰۴) مراجعه نمایید.

همان طور که چن و همکاران (۲۰۰۴) اشاره می‌کنند دقت برآوردیابی تحت روش نمونه‌گیری مجموعه رتبه‌دار در مقایسه با روش نمونه‌گیری تصادفی ساده با تعداد برابر به چگونگی رتبه‌بندی متغیر مورد علاقه بستگی دارد. در سال‌های اخیر روش‌های تعدیل یافته‌ای از روش نمونه‌گیری مجموعه رتبه‌دار با به‌کارگیری نحوه گزینش متفاوت مجموعه‌های رتبه‌دار یا نحوه رتبه‌بندی ارائه شده است. برای مثال در سال ۱۹۹۶، سماوی<sup>۲</sup> و همکاران روش نمونه‌گیری مجموعه رتبه‌دار فرین را معرفی کرده، نشان دادند که این روش در بسیاری از کاربردهای عملی برای رتبه‌بندی بسیار ساده‌تر است. در سال ۱۹۹۷ مطلق<sup>۳</sup> روش نمونه‌گیری مجموعه رتبه‌دار میان‌های را به منظور کاهش خطای رتبه‌بندی و افزایش کارایی برآوردگرها به کار برد و آن را با روش نمونه‌گیری مجموعه رتبه‌دار ساده مورد مقایسه قرار داد. وی همچنین روش نمونه‌گیری مجموعه رتبه‌دار میان‌های با استفاده از متغیر کمکی را در سال ۱۹۹۸ ارائه کرد که برای برآورد میانگین متغیر مورد نظری که رتبه‌بندی آن سخت است مفید می‌باشد. در سال ۱۹۹۹ لی<sup>۴</sup>، سینها و پرون<sup>۵</sup> بحث انتخاب‌های تصادفی را در مجموعه رتبه‌دار معرفی کرده و روش نمونه‌گیری مجموعه رتبه‌دار تصادفی (RRSS)<sup>۶</sup> را در برآورد میانگین جامعه ارائه دادند. در سال (۲۰۰۱) مطلق برآوردگرهای چندین روش نمونه‌گیری مجموعه رتبه‌دار از قبیل برآوردگر رگرسیونی مجموعه رتبه‌دار میان‌های و برآوردگر رگرسیونی مجموعه رتبه‌دار فرین در

---

<sup>1</sup> Chen

<sup>2</sup> Samavi

<sup>3</sup> Muttlak

<sup>4</sup> Li

<sup>5</sup> Perron

<sup>6</sup> Random RSS

برآورد میانگین جامعه را با هم مقایسه کرد. در سال ۲۰۰۸، جعفری جوزانی و پرون روش نمونه‌گیری مجموعه رتبه‌دار جدیدی تحت عنوان نمونه‌گیری مجموعه رتبه‌دار کامل<sup>۱</sup> FRSS را معرفی کرده و از آن در برآورد میانگین جامعه استفاده کرده و نشان دادند که برآوردگر به‌دست آمده در این روش برای میانگین جامعه بهتر از روش نمونه‌گیری تصادفی ساده است. در فصل بعد تعمیم این نتایج به برآورد سایر پارامترهای جامعه، کاربرد این روش در برآورد هسته‌ای تابع چگالی را ارائه می‌کنیم.

### ۳-۱) مفاهیم اولیه

رویکرد عمومی یک روش نمونه‌گیری مبتنی بر مجموعه‌های رتبه‌دار بر دو مرحله‌ی اساسی استوار است. در مرحله اول، نمونه‌ای تصادفی ساده از جامعه استخراج شده و با استفاده از یک مکانیسم مرتب‌سازی مشخص، بدون اندازه‌گیری صفت مورد بررسی، اعضای نمونه را بر حسب مقدار صفت و به کمک سایر منابع موجود از قبیل مقادیر متغیرهای کمکی، بررسی و قضاوت شخصی، مقادیر صفت ثبت شده در زمان‌های قبل و ... مرتب می‌کنیم. در مرحله‌ی دوم مقدار واقعی صفت مورد نظر را برای واحدی که از طریق مکانیسم مرتب‌سازی برای گزینش نهایی در نمونه انتخاب شده است اندازه‌گیری می‌کنیم.

روش مرتب‌سازی قضاوتی و سلیقه‌ای که توسط مک اینتایر مورد استفاده قرار گرفت یک نمونه از مکانیسم‌های مرتب‌سازی موجود است. امروزه از روش‌های دیگری نظیر مرتب‌سازی به کمک متغیرهای کمکی نیز استفاده می‌شود. فرض کنید مرتب‌سازی انجام شده یک مرتب‌سازی کامل باشد،

---

<sup>۱</sup> Fully RSS