دانشکده علوم ریاضی

پایان نامه

جهت اخذ درجه کارشناسی ارشد رشته آمار ریاضی

عنوان:

برآورد پارامتر بر اساس نمونه گیری مجموعه ای رتبه بندی شده

استاد راهنما :

خانم دکتر سیمیندخت براتپور

استاد مشاور:

آقای دکتر غلامرضا محتشمی برزادران

نگارش :

مریم حسنی دستگرد

بهمن ۱۳۸۸

فهرست نمادها

فارسی	انگلیسی	نماد
نمونه گیری تصادفی ساده	Simple Random Sampling	SRS
نمونه گیری مجموعه ای رتبه بنـدی شـده	Ranked Set Sampling	RSS
نمونه گیری مجموعه ای رتبه بندی شده جدید	New Ranked Set Sampling	NRSS
نمونه گیری مجموعه ای رتبه بندی اصلاح شده	Modified Ranked Set Sampling	MRSS
بهترین برآوردگر نااریب خطی	Best Linear Unbiased Estimator	BLUE
برآوردگر نااریب با مینیمم واریانس	Minimum Variance Unbiased Estimator	MVUE
برآوردگر درستنمایی ماکزیمم	Maximum Likelihood Estimator	MLE
میانگین مربعات خطا	Mean Square Error	MSE
تابع زیان مربع خطا	Square Error Loss Function	SELF

فهرست مطالب

چکیده ---------- ٦
پیش گفتار ---------- ۷
فصل اول: معرفی روش نمونه گیری مجموعه ای رتبه بندی شده ----------۹
۱ . ۱. مقدمه ---------- ۱۰
۱ . ۲. تاریخچه---------- ۱۰
۱ . ۳. مقدمه ای در پیدایش الگوی RSS---------- ۱۱
۱ . ٤. معرفی طرح RSS ---------- ۱۳
۱ .٤ .۱. معیار رتبه بندی در روش RSS ---------- ۱۸
۱ .٤ .۲. ویژگی میانگین و واریانس در روش RSS ---------- ۱۸
۱ . ٥. یادآوری مفاهیم اولیه ---------- ۲۳
۱ .٥ .۱. اصل بیز---------- ۲۳
۱ .٥ .۲. توزیع پیشین و توزیع پسین ----------.۲٤
۱ .٥ .۳. برآوردگر بیز ---------- ۲۷
۱ .٥ .٤. توابع زیان ---------- ۲۸

فصل دوم: برآورد بیزی با بکار بردن روش RSS ---------- ۳۰
۲ . ۱. مقدمه ----------.۳۱
۲ . ۲. حالت کلی برآورد بیز تحت روش RSS ---------- ۳۱
۲ .۲ .۱. رابطه بین چگالی پسین θ برحسب $\underline{x}$ و چگالی پسین θ برحسب $\underline{y}$ ---- ۳۲
۲ . ۳. خانواده نمایی با پیشین مزدوج ---------- ۳۷

۲.۳.۱. توزیع نمایی ---------- ۳۹
۲.۳.۲. مقایسه عددی ---------- ٤٨
۲.٤. کاربرد ---------- ۵۱
۲.٤.۱. خلاصه ای از داده ها و نتایج برمبنای رتبه بندی کامل ---------- ۵۱
۲.٤.۲. نتایج بر مبنای رتبه بندی بصری ---------- ۵۳

فصل سوم: روش RSS جدید برای خانواده توزیع های یک پارامتری ------- ٥٤
۳.۱. مقدمه ---------- ۵۵
۳.۲. روش NRSS برای n فرد ---------- ۵۷
۳.۳. توزیع نمایی یک پارامتری ---------- ۵۸
۳.٤. توزیع لجستیک ---------- ۷۱
۳.۵. تأثیرخطاها در رتبه بندی ---------- ۷۵
نتیجه تحقیق---------- ۷۸
ضمیمه ---------- ۷۹
واژه نامه ---------- ۸٦
منابع ---------- ۹۲

فهرست تصاویر

۱ . ۱ نمودار توزیع فراوانی قدهای رتبه دار شده در توزیع فراوانی جمعیت همه قدها ----- ۱۶

۱ . ۲ نمودار یک طرح RSS برای n = 3 ---------- ۱۷

فهرست جداول

۲ . ۱ ریسک بیز برآوردگر بیز تحت روش SRS و RSS و کارآیی ($eff = r_s / r_r$) ----- ٤۹

۲ . ۲ توابع ریسک برآوردگر بیز تعمیم یافته تحت روش SRS و RSS و کارآیی ($eff^* = R_s / R_r$). ---------- ۵۰

۲ . ۳ اریبی و MSE برآوردگر بیز تعمیم یافته و کارآیی ($\hat{eff}^* = MSE(\hat{\mu}_s) / MSE(\hat{\mu}_R)$) برای رتبه بندی کامل ---------- ۵۲

۲ . ٤ اریبی و MSE برآوردگر بیز تعمیم یافته و کارآیی ($\hat{eff}^* = MSE(\hat{\mu}_s) / MSE(\hat{\mu}_R)$) برای رتبه بندی بصری ---------- ۵۳

۳ . ۱ واریانس ها و دقت های نسبی برآوردگرهای σ برای توزیع نمایی ---------- ۷۰

۳ . ۲ آماره های مرتب، ضرایب و واریانس μ^{**} برای توزیع لجستیک ---------- ۷٤

۳ . ۳ واریانس ها و دقت های نسبی برآوردگرهای μ برای توزیع لجستیک ---------- ۷۵

۳ . ٤ دقت های نسبی شبیه سازی شده برای توزیع نمایی و لجستیک تحت روش های RSS، MRSS و NRSS ---------- ۷۷

چکیده

همان طور که می دانیم برای کنترل و ارزیابی محیط اطرافمان به داده های شهودی نیازمندیم تا بتوانیم آنها را در مقابل داده های به دست آمده از روش های دقیق آزمایشگاهی قرار داده و نتایج این دو گروه از داده ها را با هم مقایسه کنیم.

در این پایان نامه برای افزایش دقت برآوردگرها با صرف وقت و هزینه کمتر، یک روش نمونه گیری تحت عنوان نمونه گیری مجموعه ای رتبه بندی شده را معرفی کرده و بعضی از ویژگی های آن را بیان می کنیم. سپس مقایسه کارآیی برآوردگرهای بیز بدست آمده از این روش را نسبت به روش نمونه گیری تصادفی ساده به کمک ریسک بیز آن ها، مورد بررسی قرار می دهیم.

در ادامه، روش نمونه گیری جدیدی که برگرفته از روش نمونه گیری مجموعه ای رتبه بندی شده می باشد را معرفی می کنیم و به مقایسه کارآیی برآوردگرهای نااریب با مینیمم واریانس و بهترین برآوردگرهای نااریب خطی به دست آمده از این روش نسبت به روشهای نمونه گیری مجموعه ای رتبه بندی شده دیگرمی پردازیم.

واژه های کلیدی: نمونه گیری مجموعه ای رتبه بندی شده، برآوردگر بیز، ریسک بیز، کارآیی

پیش گفتار

در این پایان نامه یک روش نمونه گیری ارائه شده است که در مقایسه با روش نمونه گیری تصادفی ساده بر دقت برآوردگرهای جامعه می افزاید و از صرف هزینه و وقت زیاد می کاهد.

در فصل اول به منظور آشنایی با این روش که به نمونه گیری مجموعه ای رتبه بندی شده معروف است، تاریخچه و مقدمه کوتاهی از پیدایش آن را بیان کرده و به معرفی الگو و بعضی از ویژگی های آن می پردازیم که این مطالب در مقالات [٦] ، [٧] و [٨] آمده است. در پایان این فصل مفاهیمی را یادآور می شویم که در فهم بیشتر مطالب به ما کمک می کند.

در فصل دوم برآورد بیزی پارامتر یک توزیع با استفاده از روش نمونه گیری مجموعه ای رتبه بندی شده مورد توجه قرار می گیرد. ابتدا برآوردگر بیز را در یک حالت کلی و سپس برای خانواده نمایی، با این روش نمونه گیری بدست می آوریم. در ادامه نشان می دهیم برآوردگر بیز بر مبنای روش مورد نظر کارآتر از برآوردگر بیز بر مبنای روش نمونه گیری تصادفی ساده است، زیرا دارای ریسک بیز کوچکتری نسبت به استفاده از نمونه گیری تصادفی ساده است. این فصل بیشتر بر اساس مقالات [٧] ، [٩] و [١٥] می باشد. در پایان به یک مثال کاربردی اشاره می کنیم.

در فصل سوم روش نمونه گیری رتبه بندی شده جدیدی برای خانواده توزیع یک و دو پارامتری معرفی می شود که در مقاله [١٦] به آن پرداخته است. این روش برای n های (حجم نمونه) زوج بکار می رود در حالی که برای n های فرد جواب نمی دهد. با تغییراتی که در این فصل روی آن ها بحث می کنیم، روش معرفی شده جدید را طوری اصلاح می کنیم که بتوان برای حجم نمونه های فرد نیز آن را بکار برد.

در پایان این فصل برآوردگرهایی برای پارامترهای توزیع نمایی و لجستیک در [٢٦] بر مبنای همه روش های نمونه گیری رتبه بندی شده که در این رساله معرفی کردیم بدست می آوریم و آن ها را با هم و با بهترین برآوردگر نااریب خطی مقایسه می کنیم. در نهایت به این نتیجه می رسیم که روش نمونه گیری جدید کارآتر از روش های دیگرِ نمونه گیری مجموعه ای رتبه بندی شده و بهترین برآوردگر نااریب خطی می باشد.

در ادامه، ضمیمه و واژه نامه ارائه شده و در انتها منابعی که در پایان نامه مورد استفاده قرار گرفته اند، آمده است.

قابل ذکر است که در این پایان نامه، بخش هایی که توسط نگارنده بیان شده و یا قضایا و مطالبی که اثبات آن ها توسط نگارنده انجام شده است با علامت */ تا /* نشانه گذاری شده است.

از سرکار خانم دکتر یوسف زاده و مسئولین محترم کتابخانه جناب آقای داوود نژاد و سرکار خانم صادقی و دوست عزیزم خانم روشن نژاد به خاطر تمامی زحماتشان متشکرم.

از تمامی عزیزانی که در نگارش پایان نامه این حقیر را یاری نموده اند، تشکر و قدر دانی می نمایم.

خدایا ! به ما توفیق ده تا بتوانیم گل فرصت را که بر شاخه عمر می روید، قبل از آن که بر زمین افتد، بچینیم.

مریم حسنی دستگرد

بهمن ۸۸

فصل اول

معرفی

روش نمونه گیری مجموعه ای

رتبه بندی شده

۱.۱. مقدمه

در این فصل به بیان تاریخچه و مقدمه کوتاهی از روش نمونه گیری مورد بحث در این رساله یعنی نمونه گیریِ مجموعه ایِ رتبه بندی شده[1] (RSS) پرداخته و درباره جزئیات، چگونگی تولید یک طرح RSS و بعضی از ویژگی های آن نیز صحبت می کنیم.

در ادامه به بیان مفاهیم اولیه ای که برای درک بهتر مطالب فصل های بعدی به آنها نیاز داریم می پردازیم.

۲.۱. تاریخچه

در اوایل دهه ۱۹۵۰ ، مک اینتایر[2] [٤] یک روش نمونه گیری معرفی کرد و از آن برای برآورد میانگین محصولات مرتعی و علوفه ای استفاده کرد. او با بکارگیری این روش به صرفه جویی قابل توجهی در جمع آوری داده ها دست یافت.

بعد از مدتی در سال ۱۹٦٦، هالز و دل[3] [۵] نام نمونه گیری مجموعه ای رتبه بندی شده RSS را برای این روش در نظر گرفتند.

پس از گذشت یک دهه از طرح مک اینتایر، تحقیقات دامنه داری برای کشف تأثیر و کاربرد بیشتر RSS ادامه پیدا کرد. در اینجا به چند مورد اشاره می کنیم:

تاکاهاشی و واکیموتو[4] در مقاله [۹] ، درباره جزئیات روش RSS و کارایی آن نسبت به روش نمونه گیری تصادفی ساده[5] (SRS) مطالعه کردند.

دل و کلاتر[6] در [۱۰] نشان دادند که برآوردگر RSS برای میانگین یک جامعه، بدون در نظر گرفتن خطای رتبه بندی، نااریب است و دقت آن حداقل به اندازه برآوردگر SRS با همان تعداد حجم نمونه است.

---

1. Ranked Set Sampling
2. Mc Intyre
3. Halls and Dell
4. Takahasi and Wakimoto
5. Simple Random Sampling
6. Clutter

لوین و دیوید[۷] در مقاله [۱۱] ، رتبه بندی در روش RSS را توسط یک متغیر کمکی عددی انجام دادند.

اَل صالح و موت لاک[۸] در [۷] و [۱۲] ، برآوردگرهای بیز به روش RSS و SRS را برای برآورد پارامتر مقیاس توزیع نمایی با هم مقایسه کردند.

اَل صالح و اَل کدیری[۹] در [۱۳] ، RSS مضاعف را مورد بررسی قرار دادند.

علاوه براین، محققین زیادی درباره اینکه روش RSS برآوردگرهای دقیق تری از واریانس، تابع توزیع، پارامترهای توزیع های مختلف، پارامترهای یک مدل رگرسیون خطی ساده و... به ما می دهد، بررسی هایی انجام دادند.

۱ . ۳. مقدمه ای در پیدایش الگوی RSS

بدیهی است که ما برای کنترل و ارزیابی محیط اطرافمان، به داده های شهودی نیازمندیم، که البته می توانیم آنها را در مقابل داده های بدست آمده از آزمایشهای کنترل شده قرار دهیم و نتایج این دو گروه از داده ها را با هم مقایسه کنیم. بنابراین به منظور بدست آوردن چنین داده هایی نیازمند شناسایی واحدهای نمونه برای نشان دادن جامعه مربوطه و به دنبال آن انتخاب واحدهای خاص برای اندازه گیری مشخصه یا مشخصه های مورد علاقه هستیم.

یقیناً پرهزینه ترین بخشِ فرآیندِ بدست آوردن داده ها، تحلیل های آزمایشگاهی است در حالیکه شناسایی واحدهای نمونه در الگوی RSS موضوع نسبتاً ساده ای است.

در واقع در بسیاری از مطالعاتِ وابسته به محیط زیست[۱۰] و زیست بوم شناختی[۱۱] در اندازه گیری واقعی واحدهای نمونه ممکن است با مشکلاتی مثل صرف وقت و هزینه زیاد، کار طاقت فرسا و حتی گاهی به بارآمدن اثرات مخرب روبرو شویم. حال آنکه وقتی قادر هستیم واحدهای آزمایشی یا تجربی را

7 . Lavine and David
8 . AL_Saleh and Muttlak
9 . AL_Kadiri
10.Environmental
11.Ecological

بدون اندازه گیری واقعی متغیر مورد علاقه شناسایی و رتبه بندی کنیم، استفاده از روش RSS را خیلی مفیدتر و به صرفه تر از روش SRS برای برآورد پارامترهای جامعه می دانیم.

روش نمونه گیری تصادفی ساده، روش انتخاب n واحد از جامعه ای به حجم N است ($n < N$)، به قسمی که همه نمونه های ممکن به حجم n شانس یکسان برای انتخاب شدن داشته باشند.

مک اینتایر [٤] پس از توضیح کامل روش RSS، $\hat{\mu}_{RSS}$ را به عنوان رقیبی در مقابل $\overline{X}$، میانگین نمونه تصادفی ساده، برمبنای n مشاهده از روش SRS برای برآورد میانگین جامعه(μ) پیشنهاد کرد، از این جهت که $\hat{\mu}_{RSS}$ از $\hat{\mu}_{SRS}$ با صرفه تر و ارزانتر می باشد. از اینرو مقایسه این روش نمونه گیری و روش SRS با حجم نمونه مساوی n، ممکن می باشد.

خوشبختانه در بسیاری از مطالعات آزمایشی بخصوص در کشاورزی، شناسایی و رتبه بندی واحدهای نمونه گیری با روش RSS، بدون اندازه گیری واقعی متغیر مورد علاقه، نسبتاً ارزان می باشد. بنابراین اگر قادر باشیم با استفاده از این روش، تعداد زیادی از واحدهای نمونه را برای نشان دادن جامعه مورد علاقه شناسایی کنیم، می توانیم به صرفه جویی قابل توجهی دست یابیم.

برای درک این موضوع به مفهوم پایه ای روش RSS و ارتباط آن با SRS به طور مختصر اشاره می کنیم:

فرض کنید $X_1, X_2, \ldots, X_n$ یک نمونه تصادفی از تابع توزیع $F(x)$، با میانگین μ و واریانس متناهی σ^2 باشد. به سادگی می توان نشان داد ([۱] را ببینید) که برآوردگر نااریب μ (میانگین جامعه) در روش SRS عبارتست از:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad .$$

که واریانس آن $\frac{\sigma^2}{n}$ است.

در مقایسه با روش SRS ، نمونه روش RSS تشکیل می شود از: کوچکترین مشاهده از اولین نمونه تصادفی بدست آمده به حجم n، یعنی $X_{1:n} = X_{(11)}$ به عنوان اولین عضو نمونه RSS ، دومین مشاهده کوچک از نمونه تصادفی دوم با n مشاهده که مستقل از نمونه اول است، یعنی $X_{2:n} = X_{(22)}$

به عنوان دومین عضو و... در نهایت بزرگترین مشاهده از نمونه آخر با n مشاهده که مستقل از نمونه های قبلی است، یعنی $X_{n:n} = X_{(nn)}$ به عنوان آخرین عضو ، که همه این مشاهدات بطور بصری و بدون اندازه گیری های واقعی مشخص شده اند.

اکنون به چند مثال از موارد کاربردی RSS که در مقاله [۷] و [۹] آمده است اشاره می کنیم؛

مثال ۱. ۱: فرض کنید متغیر مورد علاقه، طول نوعی از سلولهای باکتریایی در یک کشت میکروسکوپی[۱۲] است که می توان آن را با یک میکرومتر اندازه گیری کرد، که در اکثر موارد می توان ترتیب بزرگی دو یا سه سلول را در یک کشت میکروسکوپی مشابه با یک نظر اجمالی پیدا کرد.

مثال ۱. ۲: فرض کنید متغیر مورد علاقه تعداد نوعی از سلولهای باکتریایی در یک واحد حجم باشد. اگر تعدادی لوله آزمایش را که حاوی سوسپانسیون[۱۳] سلول هستند، در دست داشته باشیم، این لوله ها را می توان براساس ترتیب تجمع بیشترِ سلولها در لوله ها با استفاده از وسایل نوری مثل میکروسکوپ نوری، بدون پیدا کردن تعداد واقعی سلولهای باکتری مرتب کرد.

مثال ۱. ۳: اگر متغیر مورد علاقه را قد درختان یک منطقه خاص درنظر بگیریم. در این مورد می توانیم با یک نظر اجمالی دو یا سه درخت که نزدیک به یکدیگر هستند را مرتب کنیم.

مثال ۱. ٤: در یک مثال دیگر فرض می کنیم که علاقمند به برآورد متوسط محصول شیر گوسفندان در یک منطقه مشخص هستیم. در این مورد محصول شیر دو، سه یا چهار گوسفند به آسانی توسط صاحب گوسفندان یا توسط هر شخص، بر مبنای حجم پستان آنها رتبه بندی می شود.

۱.۴. معرفی طرح RSS

برای تولید مجموعه های رتبه بندی شده که نمونه RSS را از آن ها بدست می آوریم به ترتیب زیر عمل می کنیم:

مرحله اول : n^2 واحد نمونه اولیه را به طور تصادفی از جامعه انتخاب می کنیم.

12. Microscopic field
13. Suspension

مرحله دوم : n^2 واحد انتخاب شده مرحله اول را بطور تصادفی به n مجموعه، هریک با حجم n اختصاص می دهیم.

مرحله سوم : در این مرحله واحدهای هر مجموعه را در درون خود آن مجموعه به طور بصری یا براساس استنباطی که از مقادیر وابسته به متغیر مورد علاقه داریم، رتبه بندی می کنیم و چون این کار بطور تخمینی صورت می گیرد، حجم مجموعه ها را به خاطر می نیمم کردن خطای رتبه بندی کوچک می گیریم. بنابراین n^2 مشاهده مرتب شده در n نمونه به شکل زیر ظاهر می شود :

$$\begin{matrix} x_{(11)} & \cdots & x_{(12)} & x_{(1n)} \\ x_{(21)} & \cdots & x_{(22)} & x_{(2n)} \\ \vdots & \ddots & \vdots & \vdots \\ x_{(n1)} & \cdots & x_{(n2)} & x_{(nn)} \end{matrix}$$

به علت اهمیت موضوع رتبه بندی در این مرحله، در بخش (۱.٤.۱) به آن می پردازیم.

مرحله چهارم : در این مرحله نمونه اصلی برای تحلیل و بررسی را انتخاب می کنیم. این نمونه شامل :

۱. کوچکترین واحد رتبه بندی شده از مجموعه اول یعنی $x_{1:n} = x_{(11)}$.

۲. دومین واحد کوچک رتبه بندی شده از مجموعه دوم یعنی $x_{2:n} = x_{(22)}$ و به همین ترتیب بزرگترین واحد رتبه بندی شده از مجموعه nام است.

نکته مهم این است که اگر چه در روش RSS به تشخیص n^2 عنصر (n نمونه ی n عضوی) نیاز داریم اما فقط n تای آنها یعنی $x_{(11)}, x_{(22)}, \ldots, x_{(nn)}$ را به طور واقعی و دقیق اندازه گیری می کنیم. این مجموعه « نمونه ی مجموعه ایِ رتبه بندی شده » یعنی RSS نامیده می شود. واضح است که اعضای تشکیل دهنده RSS مستقل هستند اما هم توزیع نیستند. */ زیرا هر یک از n عضو نمونه نهایی از یک نمونه تصادفی که مستقل از دیگر نمونه ها بوده بدست آمده است و علاوه براین $x_{(ii)}$ مانند $x_{i:n}$، که iامین آماره مرتب در یک نمونه تصادفی به حجم n از توزیع $F(x)$ است، توزیع شده است و می دانیم که توزیع احتمالی آماره های مرتب نیز متفاوت است /* .

مرحله پنجم : مراحل ۱ تا ٤ را m بار تکرار می کنیم تا حجم نمونه RSS دلخواه را با توجه به رابطه زیر برای تجزیه و تحلیل بدست آوریم.

حجم نمونه دلخواه

$$r = m \cdot n \rightarrow \text{حجم یک نمونه RSS تا پایان مرحله ٤} \qquad (۱.٤.۱)$$

تعداد چرخه ها

در بعضی مقالات، به این نمونه، BRSS، « نمونه مجموعه ای رتبه بندی شده متعادل »[14] [۷] یا RSS با تخصیص برابر می گویند [٦] ، زیرا تعداد نمونه های اختصاص داده شده به هریک از چرخه ها و تعداد عضوهای انتخاب شده از تک تک آنها با هم برابر است.

برای بیان ساده ای از مفاهیم گفته شده به مثال زیر توجه کنید([٦] را ببینید):

مثال ۱. ۵: فرض کنید می خواهیم از یک نمونه تصادفی ساده با سه دانشجو، میانگین قد دانشجویان یک دانشگاه را برآورد کنیم. علاوه بر آن به منظور از بین بردن شک و تردید باید این برآورد را با یک فاصله اطمینان برای میانگین واقعی جامعه که مطابق با انتظار ماست، نشان دهیم.

روش اول : ساده ترین راه برای بدست آوردن نمونه مورد نظر این است که سه دانشجو از جامعه دانشگاه را بطور تصادفی انتخاب کنیم، سپس قد آنها را اندازه بگیریم.

نتیجه : علی رغم اینکه میانگین حسابی این سه قد یک برآورد نقطه ای نااریب از میانگین جامعه است، بازه اطمینان مربوطه می تواند خیلی بزرگ باشد که نشان دهنده اطمینان پایین و شک و تردید بالاست. زیرا میانگین یک جامعه بزرگ را تنها با سه اندازه برآورد کرده ایم. ازطرفی ما کنترلی روی افراد جامعه که وارد نمونه می شوند نداریم. برای مثال ممکن است بطور اتفاقی دو فرد خیلی کوتاه و یک فرد خیلی بلند و یا سه فرد خیلی بلند را انتخاب کرده باشیم. تنها راه غلبه بر چنین مشکلی در روش SRS این است که حجم نمونه را افزایش دهیم، که این کار هزینه زیادی را صرف می کند.

روش دوم : از طرف دیگر ما می توانیم یک نمونه به روش RSS نیز بدست آوریم. برای انجام این کار به ترتیب زیر عمل می کنیم :

14. Balanced Ranked Set Sampling

۱. ابتدا بطور تصادفی سه دانشجو را به وعده صبحانه دعوت می کنیم و بطور بصری آنها را به ترتیب قد رتبه بندی می کنیم. سپس دانشجویی را که عقیده داریم کوتاه ترین است، انتخاب کرده و قد او را بطور دقیق اندازه می گیریم.

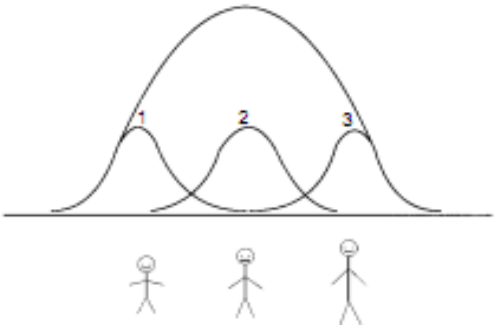
۲. این کار را برای ناهار هم انجام می دهیم و شخصی را انتخاب می کنیم که در رتبه بندی ما قد متوسط دارد.

۳. همچنین برای شام، بلندترین فرد رتبه بندی شده را انتخاب می کنیم و قدش را اندازه می گیریم.

اندازه های بدست آمده از قدِ سه فرد انتخاب شده یعنی $x_{(11)}, x_{(22)}, x_{(33)}$ یک نمونه مجموعه ای رتبه بندی شده یا RSS را تشکیل می دهد.

نتیجه: میانگین اندازه های RSS نیز مانند اندازه های نمونه SRS یک برآورد نااریب از میانگین جامعه را می دهد که به اثبات آن در بخش (۱ .٤ .۲) اشاره خواهد شد. به هرحال فاصله اطمینان مربوطه می تواند خیلی دقیق تر از فاصله ای باشد که با اندازه های SRS بدست می آید. زیرا اندازه های بدست آمده از روش RSS خیلی منظم تر از اندازه های SRS انتخاب و فاصله بندی می شوند، بنابراین نماینده بهتری برای میانگین جامعه هستند.

درواقع ما بطور تصادفی از زیر جامعه هایی که شامل دانشجویان کوتاه، متوسط و بلند هستند یک نمونه گرفته ایم، بدون اینکه زیرجامعه ها را طبقه بندی کنیم. با توجه به شکل (۱ .۱) می بینیم که چطور جامعه اصلی به زیرجامعه ها افراز شده است.



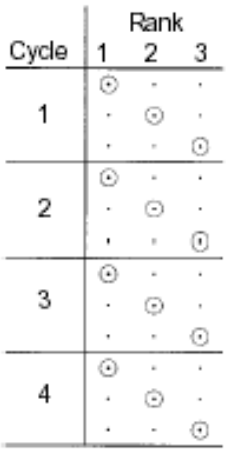
شکل ۱ .۱: توزیع های فراوانی قدهای مختلف رتبه دار شده در توزیع فراوانی جمعیت همه قدها.

اکنون اگر بخواهیم در مثال قبل حجم نمونه نهایی ۱۲ نفر باشد، طبق رابطه (۱ .٤ .۱) داریم :

$$\Rightarrow 12 = m \cdot 3 \ , \ m = 4$$

پس باید مراحلی را که برای انتخاب $n = 3$ نفر در یک روز انجام دادیم، $m = 4$ بار تکرار کنیم.

این حالت در شکل (۱ .۲) نمایش داده شده است.



شکل ۱ .۲: یک طرح RSS با حجم نمونه $n = 3$، تعداد چرخه های نمونه گیری $m = 4$ و هر سطر به یک نمونه مرتب شده بصری از افرادِ دعوت شده به یک وعده روزانه اختصاص داده شده که واحد های نمونه اصلی با دایره مشخص شده اند.

همانطور که گفته شد اگر همین تعداد دانشجو را به روش SRS انتخاب کنیم، ممکن است همه ۱۲ واحد از افراد نسبتاً کوتاه یا خیلی بلند آمده باشند در حالیکه فقط دو یا سه فرد متوسط انتخاب شده باشند. بنابراین تنها راه افزایش احتمال پوشش دادن همه دانشجویان در انتخاب نمونه SRS این است که حجم نمونه را افزایش دهیم، درحالیکه با روش RSS با همان تعداد واحد نمونه می توان نماینده بهتری برای جامعه بدست آورد. بنابراین در هزینه های اندازه گیری به طور چشم گیری صرفه جویی خواهیم کرد.

۱.۴.۱. معیار رتبه بندی در روش RSS

راهنمای اصلی برای موفقیت در اجرای طرح RSS در مرحله سوم که همان مرحله رتبه بندی نمونه هاست، قرار دارد. که این رتبه بندی ممکن است به چند طریق انجام شود، در یکی از آنها رتبه بندی براساس قضاوت شخصی یعنی بازبینی بصری یا نظر کارشناس درباره واحدهای نمونه صورت می گیرد.

مثال ۱. ٦ : یک جنگل بان به آسانی می تواند از لحاظ حاصلخیزی، سه یا چهار قطعه زمین زراعی را با لحاظ کردن حجم یا بزرگی توده علفها رتبه بندی کند.

یکی دیگر از راههای رتبه بندی این است که اگر مشخصه دیگری در دسترس باشد که با مشخصه موردعلاقه همبستگی زیادی داشته ولی هزینه های بدست آوردنش خیلی کمتر باشد، آنگاه می توانیم رتبه بندی را توسط مقادیر چنین متغیر کمکی انجام دهیم.

مثال ۱. ۷ : یک کارشناس زباله های خطرناک به آسانی می تواند مناطقی از خاک را با در نظر گرفتن آلوده کننده های سمی، براساس شواهد ظاهری مثل سطوح لکه لکه، بی رنگ شدن یا ظاهر تحت فشار میوه ها و سبزیجات رتبه بندی کند.

مثال ۱. ۸ : در مثالی دیگر می بینیم شدت بازتاب امواج الکترو مغناطیس نزدیک مادون قرمز که در عکسهای دیجیتالی کوچک و حساس ثبت می شوند به طور مستقیم با تمرکز سبزیجات روی زمین تناسب دارند. بنابراین به جای اینکه با متغیر اصلی شدت بازتاب امواج کار کنیم، می توانیم با متغیر کمکی همبسته با آن، رتبه بندی را انجام دهیم.

۲.۴.۱. ویژگی میانگین و واریانس در روش RSS

فرض می کنیم $x_n,\ldots,x_1$ یک نمونه SRS با حجم n از یک جامعه پیوسته با تابع چگالی[۱۵] $f(x)$ و تابع توزیع[۱٦] $F(x)$ باشد. اگر $x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$ آماره های مرتب وابسته به مشاهدات

15. Density function
16. Distribution function