سیزدهمین کنفرانس آمار ایران
۴-۲ شهریور ۱۳۹۵

طرح نمونه‌گیری فضایی متعادل دومرحله‌ای برای پیش‌گویی

رامین خاورزاده[۱]، محسن محمدزاده [۲]

[۱] دانشجوی دکتری آمار، دانشگاه تربیت مدرس

[۲] گروه آمار، دانشگاه تربیت مدرس

چکیده: اغلب در بررسی‌های نمونه‌ای فرض بر آن است که اعضای نمونه، از جامعه‌ای با واحدهای مستقل گرفته‌شده است. این فرض در تمامی مراحل نمونه‌گیری تحلیل و مدل‌سازی مورداستفاده قرار می‌گیرد. اما وقتی اعضای جامعه مورد مطالعه به‌نوعی وابسته باشند، تمامی مراحل آماری و حتی روش‌های نمونه‌گیری نیازمند بازنگری و لحاظ کردن ساختار همبستگی داده‌ها خواهند بود. از طرفی در نمونه‌گیری کلاسیک، چنانچه متغیرهای کمکی وجود داشته باشند، برای ارتقاء کیفیت طرح نمونه‌گیری می‌توان از نمونه‌گیری متعادل استفاده کرد. در این مقاله نمونه‌گیری فضایی متعادل دو مرحله‌ای معرفی می‌شود که در آن از مؤلفه‌های موقعیت‌های فضایی به‌عنوان متغیرهای کمکی برای برقراری تعادل استفاده شده است. سپس با استفاده از ملاک ریشه توان دوم خطای پیش‌گویی فضایی نشان داده می‌شود که پیش‌گویی بر اساس نمونه‌گیری فضایی متعادل نسبت به روش‌های دیگر نمونه‌گیری متحمل خطای کمتری می‌شود. بعلاوه نشان داده می‌شود نمونه‌گیری فضایی متعادل بیشترین پوشش بر ناحیه‌ی مورد مطالعه را فراهم می‌سازد. در ادامه کارایی این روش‌های مختلف نمونه‌گیری در مسئله کاهش حجم ذخیره‌ی عکس‌ها بر اساس پیش‌گویی فضایی از پیکس‌های نمونه‌گیری شده مورد بررسی قرار می‌گیرند.

واژه‌های کلیدی: نمونه‌گیری فضایی، نمونه‌گیری متعادل، نمونه‌گیری فضایی بهینه.

۱ مقدمه

برخلاف روش‌های معمول آمار که برای تحلیل و نمونه‌گیری از فرض استقلال مشاهدات استفاده می‌شود، در آمار فضایی، داده‌هایی مورد تحلیل قرار می‌گیرند که همبسته بوده و این وابستگی ناشی از موقعیت قرارگیری آن‌ها در فضای موردمطالعه است که همبستگی فضایی نامیده می‌شود. مدل بندی داده‌های فضایی به‌طور معمول توسط میدان تصادفی $\{z(s);\ s \in D\}$ انجام می‌شود، که در آن مجموعهٔ اندیس گذارD زیرمجموعه‌ای از فضای اقلیدسی d بعدی R^d ، $d \geq ۱$ است.

[۲] رامین خاورزاده : r.khavarzade@modares.ac.ir

فرض کنید موقعیت‌های فضایی اعضای جامعه موردنظر به‌صورت $U = \{s_1, ..., s_N\}$ نشان داده شوند و هدف پیش‌گویی مقدار میدان تصادفی در یک موقعیت دلخواه با استفاده از مشاهدات در نقاط نمونه‌گیری $\{s_1, ..., s_N\}$ باشد. ازآنجا که محل قرارگیری اعضای نمونه انتخاب‌شده، بر عملکرد پیش‌گو کاملاً مؤثر است، انتخاب نمونه مناسب برای پیش‌گویی از اهداف نمونه‌گیری فضایی به‌حساب می‌آید.

طرح نمونه‌گیری برای بررسی‌های نمونه‌ای مستقل به‌طور جامع در متون آماری موردبحث و بررسی قرارگرفته‌اند. کوکران(۱۹۳۶)، ساراندال و همکاران(۱۹۹۲)، تیله (۲۰۰۶) و ریموند چامبرز و رابرت کلارک (۲۰۱۲) شرح مبسوطی از مشخصه‌های طرح‌های نمونه‌گیری را ارائه کرده‌اند که شامل تعاریف روشنی از جامعه موردمطالعه، واحدهای نمونه‌گیری، چارچوب نمونه‌گیری و چگونگی انتخاب نمونه است. طرح‌های مرتبط با پژوهش‌ها و سنجش‌های محیطی که در آن‌ها واحدهای جامعه لزوماً از یکدیگر مستقل نیستند، روش‌های نمونه‌گیری متداول را با چالش‌های جدیدی روبه‌رو کرده است. نمونه‌گیری از منابع محیطی به سیستم‌های پیچیده‌ای نیاز دارد. که در آن ممکن است با زمان یا مکان، حرکت یا تغییر کنند. به‌علاوه همواره ساختن چارچوب نمونه‌گیری قابل‌اعتماد، از جامعه هدف به‌سادگی میسر نیست. حتی ممکن است برخلاف روش‌های نمونه‌گیری معمول، چارچوب مشخصی از واحدهای جامعه مورد نظر، قابل‌تعریف نبوده و تنها، واحدهای جامعه بر اساس مکان قرارگیری آن‌ها در محیط قابل‌تعریف باشند. در این صورت با توجه به وجود یک جامعه گسترده، تنها قادر به نمونه‌گیری بخش کوچکی از جامعه خواهیم بود. یکی از مهم‌ترین مسائلی که در چنین جوامعی باید به آن توجه شود، مشابهت واحدهای نزدیک‌تر ازلحاظ مکانی است. بنابراین لازم است مختصات موقعیت‌های فضایی واحدها در تحلیل نمونه‌ای لحاظ شود. ازاین‌رو برای چنین جوامعی روش‌های جدیدی تحت عنوان روش‌های نمونه‌گیری فضایی معرفی شده است.

در اولین مطالعات انجام‌شده درزمینه‌ی نمونه‌گیری فضایی می‌توان به کار مک بارانتی و همکاران (۱۹۸۱) اشاره کرد. آن‌ها نشان دادند اگر هدف، پیش‌گویی و تعیین پهنه‌بندی فضایی متغیر مورد مطالعه با استفاده از کریگیدن[1] ، بدون در نظر گرفتن اثرات مرزی باشد به کمک مشبکهٔ مثلثی از موقعیت‌های نمونه‌ای، واریانس خطای پیش‌گویی تحت فرض‌های مانایی و همسانگردی مینیمم می‌شود. در این میان طرح‌های نمونه‌گیری دیگری نیز با رویکرد طرح مبنا معرفی‌شده‌اند که بین آن‌ها می‌توان به طرح نمونه‌گیری موزون بارد واحدهای مجاور (هدایت و همکاران، ۱۹۸۸) و روش دنباله‌ای واحدهای ناحیه‌ای وابسته (آرابیا، ۱۹۹۳) و طرح‌های طبقه‌ای موزاییک بندی تصادفی تعمیم‌یافته (اتیونز و السن ۲۰۰۴) اشاره کرد. اما هر یک از این روش‌های با محدودیت‌های مختلفی روبرو هستند. به‌عنوان‌مثال، روش ”نمونه‌گیری موزون بارد واحدهای مجاور“ برای گرفتن نمونه از فضایی یک‌بعدی در نظر گرفته می‌شود. این روش تنها از قرار گرفتن واحدهای مجاور در نمونه ممانعت می‌کند. در روش ”دنباله‌ای واحدهای ناحیه‌ای وابسته“ نمونه‌ها به‌صورت دوبعدی در نظر گرفته می‌شوند، اما از آنجا که ناحیه‌های در نظر گرفته شده خصوصاً در قسمت‌های مرزی، تعداد واحدهای مجاور برابر ندارند، در عمل محاسبات با مشکل مواجه می‌شوند. بعلاوه درروش ”طرح‌های طبقه‌ای موزاییک بندی تصادفی تعمیم‌یافته“ که تعمیمی از دو روش،تحت عناوین ”طرح نمونه‌گیری طبقه‌ای موزاییک بندی“ و ”طرح طبقه‌ای موزاییک بندی تصادفی“ است، تنها الگویی برای انتخاب مکان اعضای نمونه معرفی‌شده است که دارای بیشترین پوشش در ناحیهٔ مورد نظر باشند. بیشترین تمرکز مطالعات در نمونه‌گیری فضایی در سال‌های

[1] Kriging

اخیر بر روی ساختارهای هندسی باهدف تولید یک مشبکِ تصادفی و بهینه‌سازی نسبت به اندازه‌ی نمونه، واریانس و خودهمبستگی فضایی داده‌ها بوده است (دسراد و بارهن، ۲۰۰۵؛ صالحی، ۲۰۰۴).

دیگرویجتر و همکاران (۲۰۰۶) و دوبی و همکاران (۲۰۰۸) روش‌های مختلف نمونه‌گیری را برای داده‌های فضایی مورد استفاده قراردادند و نحوه انجام، مزیت‌ها و معایب این طرح‌ها را به تفضیل بیان کردند. به هرحال در روش‌های معرفی‌شده همچنان رویکرد نمونه‌گیری از همان اصول روش‌های کلاسیک نمونه‌گیری فضایی بوده است.

روش‌های مختلف نمونه‌گیری فضایی به‌منظور اهداف و مسائل گوناگون مطرح‌شده‌اند. بخشی از آن‌ها می‌تواند همان اهداف نمونه‌گیری کلاسیک، یعنی برآورد پارامترهای جامعه هدف، مانند میانگین، مقدار کل یا نسبت باشد. پارامترهایی از این قبیل را می‌توان با دقتی معمول از نمونه احتمالاتی به کمک آنچه بدان رویکرد "طرح مبنا" گفته می‌شود، برآورد کرد. از طرف دیگر هاینینگ (۱۹۹۰) نوعی نمونه‌گیری را برای بررسی ساختار همبستگی بر اساس تغییرنگار یا هم‌تغییرنگار، مطرح کرده است، که بخش دیگر اهداف نمونه‌گیری فضایی را تشکیل می‌دهد. هدف دیگر مطالعه، علاوه بر پارامترهای جامعه، می‌تواند پیش‌گویی فضایی باشد. نمونه‌گیری فضایی با طراحی شبکه‌های سنجش نیز در ارتباط است. به عنوان مثال برای تغییر محل ایستگاه‌های سنجش آلودگی هوا، باران‌سنجی، سنجش سطوح ازون در یک شهر، مکان‌یابی پارک‌ها، ایستگاه‌های آتش‌نشانی یا توزیع فضایی ایستگاه‌های پلیس برای امنیت بیشتر شهر، از مدلی آماری استفاده می‌شود. درواقع در چنین رویکرد استنباطی که بدان "مدل مبنا" گفته می‌شود، هدف تعیین طرح بهینه به‌منظور پیش‌گویی در موقعیت‌ها یا نواحی فاقد مشاهده، برآورد پارامترهای تابع کوواریانس فضایی، برآورد ضرایب رگرسیونی روند یا پوشش مناسب نمونه‌ای است. بنابراین اهداف مطالعه نقش بسزایی در تعیین طرح نمونه‌گیری مناسب دارند، که ممکن است در مسئله دیگر کاملاً متفاوت باشد. در این مقاله از مفهوم نمونه متعادل[۲] در نمونه‌گیری کلاسیک، به نمونه‌گیری فضایی متعادل[۳] ، برای به دست اوردن نمونه‌ی بهینه تعمیم داده می‌شود. برای دستیابی به نمونه متعادل از تکنیک روش مکعبی استفاده شده است و در انتها روش نمونه‌گیری فضایی متعادل با نمونه‌گیری تصادفی ساده بر اساس ملاک میانگین توان دوم خطای کریگیدن مقایسه می‌شود.

۲ نمونه‌گیری متعادل

در نمونه‌گیری کلاسیک برای متغیر پاسخ Y، چنانچه یک متغیر کمکی مانند بردار تصادفی Z وجود داشته باشد، یتس (۱۹۴۶) برای ارتقاء کیفیت طرح نمونه‌گیری، نمونه متعادل را معرفی کرد، که در آن نمونه‌ای از تحققهای متغیر پاسخ Y بر روی متغیر کمکی Z، متعادل نامیده می‌شود هرگاه در نمونه مقادیر به‌گونه‌ایz انتخاب شوند که میانگین نمونه‌ای آن‌ها دقیقاً برابر مقدار واقعی میانگین متغیر Z باشد. رویال و هرسون (۱۹۷۳) شرط قوی‌تر انطباق گشتاورهای اول نمونه‌ای Z با گشتاورهای جامعه متناظرشان را مطرح کردند. بینش نهفته در پس این متعادل‌سازی، این است که، با انطباق گشتاورهای نمونه‌ایZ با گشتاورهای جامعه، تعادلی تقریبی روی Y ایجاد شود، تا نمونه متعادل انتخابی یک نمونه معرف بهتر از نمونه‌های مستخرج از جامعه هدف باشد. درواقع نمونه‌ای معرف نامیده می‌شود که منعکس‌کننده جامعه در تمام ابعاد مورد نظر باشد به‌نوعی که نتایج حاصل از تحلیل مشاهدات نمونه با نتایج حاصل از کل جامعه تقریبا همسان شوند (هایک، ۱۹۸۱). با این تعریف می‌توان با متعادل کردن نمونه، به یک نمونه معرف دست یافت.

فرض کنید جامعه مورد نظر شامل N واحد آماری به‌صورت (۰.۲)$U = \{1, ..., N\}$باشد، که اطلاعات آن معمولاً مقادیر تجمعی متغیرهای مورد بررسی، به‌صورت $t_y = (y_1 + ... + y_N) = \sum_{k\in U} y_k$هستند. در نمونه‌گیری احتمالاتی، که انتخاب نمونه بر

[۲] Balanced sample

[۳] Spatial balanced sampling

اساس احتمالات انجام می‌شود، احتمال انتخاب واحد u_kاز جامعه به‌عنوان عضو نمونه عدد مشخص مثبت π_kاست، که احتمال شمول u_kنامیده می‌شود. بدیهی است مجموع احتمالات شمول برای کلیه واحدهای جامعه برابر با اندازه‌ی نمونه انتخابی است.

۱.۲ نمونه‌گیری متعادل با روش مکعبی

روش مکعبی یکی از روش‌های مرسوم برای دست‌یابی به نمونه متعادل، با چند متغیر کمکی و با احتمالات شمول متفاوت است. این روش شامل دو مرحله فراز و فرود است. در مرحله‌ی فراز قیدهای تعادل همواره برقرار هستند، به این صورت که مجموعه‌ای از نمونه‌ها در نظر گرفته می‌شود که همگی آن‌ها شرط تعادل را برقرار می‌کنند. اما در این نمونه‌ها اعضای جامعه به‌صورت کامل در نمونه قرار نمی‌گیرند. به‌عنوان مثال، چنانچه هدف انتخاب یک نمونه با اندازه ۲ از یک جامعه با ۴ عضو باشد. ممکن است یکی از نمونه‌ها در مرحله فراز به‌صورت نیمی از تمامی اعضای جامعه باشد، یعنی $\frac{۱}{۲}$از واحد اول و $\frac{۱}{۲}$از واحد دوم و $\frac{۱}{۲}$از واحد سوم و $\frac{۱}{۲}$از واحد چهارم که شرط تعادل را محقق می‌سازد. سپس با روشی کاملاً تصادفی یکی از نمونه‌ها که شرط تعادل را محقق می‌سازد، انتخاب می‌شود. در این مرحله چنانچه تمامی واحدهای نمونه انتخاب شده به‌صورت کامل باشند، نمونه متعادل مورد نظر انتخاب شده و نمونه‌گیری به پایان خواهد رسید. اما چنانچه نمونه شامل واحدی باشد که به‌صورت کامل در نمونه قرار نگرفته باشد، الگوریتم نمونه‌گیری وارد مرحله‌ی بعد می‌شود. در مرحله‌ی بعدی، یعنی مرحله فرود، با توجه به احتمال شمول هر یک از اعضای جامعه اعضای نمونه‌ای که به‌صورت کسری در نمونه قرارگرفته‌اند به یکی از مقادیر یک (یعنی حضور در نمونه) یا صفر (یعنی عدم حضور این عضو در نمونه)تبدیل می‌شود. هدف این است که برای همه احتمال شمول‌ها این کسرها به‌طور تصادفی به ۰ یا ۱ گرد شوند تا درنهایت به‌اندازه تعداد حجم نمونه، عضو کامل به دست آید. مرحله‌ی فرود لزوماً به برقراری کامل شرط تعادل منجر نمی‌شود بلکه تنها تعادلی نسبی روی متغیرهای کمکی ایجاد می‌کند.

۲.۲ روش نمونه‌گیری متعادل دومرحله‌ای

چنانچه با رویکردی فضایی به طرح نمونه‌گیری متعادل نگاه شود و موقعیت‌های آماری به‌صورت $U = \{s_۱, ..., s_N\}$در نظر گرفته شوند، شرط تعادل فضایی برای گشتاور اول به‌صورت $\sum_{s_k \in U} X_{s_k} = \sum_{s_k \in U} \frac{x_{s_k} c_{s_k}}{\pi_{s_k}}$و برای گشتاورm ام نیز به صورت $\sum_{s_k \in U} {X^m}_{s_k} = \sum_{s_k \in U} \frac{{x^m}_{s_k} c_{s_k}}{\pi_{s_k}}$تعریف می‌شوند.

اگر متغیرهای کمکی $x_۱$ طول جغرافیایی، $x_۲$ عرض جغرافیایی و $x_۳$ ارتفاع واحدهای جامعه در نظر گرفته شوند، می‌توان با متعادل کردن بر اساس این متغیرهای کمکی به نمونه‌های بهینه متعادل رسید. وقتی موقعیت‌های فضایی به عنوان متغیرهای کمکی در نظر گرفته شوند، نمونه متعادل فضایی نامیده می‌شود که گشتاورهای فضایی مختصات موقعیت‌های نمونه‌ای بر گشتاورهای فضایی جامعه منطبق باشند. گشتاورهای فضایی مرتبه اول (گرانیگاه) و گشتاور فضایی مرتبه دوم (اینرسی) مشابه با ماتریس کوواریانس است و نظم شکل ناحیه مورد مطالعه یا نظم الگوی نقاط نمونه‌ای را اندازه می‌گیرند (دوبی و همکاران ۲۰۰۸). نمونه‌گیری متعادل فضایی، عناصر نمونه‌گیری سیستماتیک و تصادفی ساده را با یکدیگر ترکیب می‌کند. در این روش موقعیت‌ها به‌طور تصادفی انتخاب می‌شوند به‌گونه‌ای که پوشش کامل روی جامعه را تضمین می‌نماید.

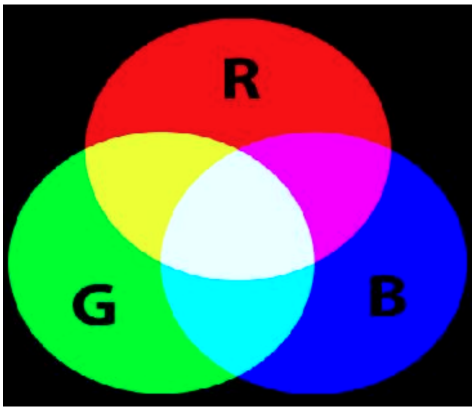
روش نمونه‌گیری متعادل متضمن نمونه‌ای است که به‌صورت کاملاً متعادل بر روی متغیرهای کمکی پخش‌شده باشد. حال چنانچه عرض و طول مختصات فضایی به عنوان متغیر کمکی در نظر گرفته شوند، نمونه حاصل، نمونه متعادل شده بر روی این متغیرهای کمکی به‌صورت کاملاً متعادل انتخاب‌شده‌اند اما لزوماً در کل ناحیه موردبررسی پخش نشده‌اند. برای رفع این مشکل، روش نمونه‌گیری متعادل دومرحله‌ای را معرفی می‌کنیم.

در مرحله اول چارچوب فضایی با توجه به حجم نمونه طبقه‌بندی می‌شود به گونه‌ای که فضایی نمونه‌ای به چند برابر (۵ تا ۱۰ برابر) حجم نمونه تقسیم‌بندی می‌شود. طبیعتاً هرچه حجم نمونه بیشتر شود، طبقه‌بندی در مرحله اول بیشتر و امکان حصول نمونه متعادل اولیه بیشتر خواهد بود. در مرحله دوم به تعداد اندازه نمونه طبقات با روش مکعبی انتخاب می‌شوند. این امر باعث می‌شود که طبقات نسبت به طول و عرض جغرافیایی متعادل باشند. سپس عضوی در هر طبقه انتخاب‌شده به‌عنوان نمونه انتخاب می‌شود که فاصله‌ی بیشتری را از طبقات انتخاب‌شده همسایه خود داشته باشد. به‌این‌ترتیب نمونه‌های نهایی، هم به لحاظ عرض و طول جغرافیایی متعادل هستند و هم بر روی فضای دوبعدی کاملاً پخش‌شده است. از طرف دیگر چون نمونه‌هایی که در ناحیه‌ی دوبعدی انتخاب می‌شوند در مرحله دوم بیشتر به سمت حاشیه‌ها رانده می‌شوند، نمونه‌های حاشیه‌ای تمایل بیشتری برای رفتن به حاشیه فضای دوبعدی خواهند داشت.

۳ مثال کاربردی

مدل رنگ RGB برای ایجاد تصاویر در تلویزیون و مانیتورها به کار گرفته می‌شود، که در آن تمام‌رنگ‌ها از ترکیب سه رنگ قرمز(R)، سبز (G) و آبی (B) تشکیل می‌شوند. با ترکیب رنگ‌های ابتدایی[۴] ، رنگ‌های دیگر یا ثانویه[۵] ایجاد می‌شوند. سیستم افزودنی[۶] از ترکیب سه رنگ اصلی RGB رنگ سفید حاصل می‌شود. با توجه به اینکه صفحه‌نمایش این قبیل وسایل تیره است برای ایجاد تصویر باید به آن رنگ اضافه شود.

یکی از کاربردهای نمونه‌گیری فضایی، کاهش حجم تصاویر است. در شکل ۲ تصویر برج آزادی در شهر تهران نشان داده شده است. این عکس متشکل از ۲۲۶۲ * ۳۲۴۹ پیکسل است، که هر یک از سه رنگ آبی، قرمز و سبز تشکیل شده‌اند.



شکل ۱: نمایش رنگ‌ها به‌وسیله سه رنگ قرمز، آبی و سبز.

وجود همبستگی فضایی در این سه تفکیک عکس کاملاً مشهود است اما جدای از همبستگی فضایی این سه رنگ ازلحاظ مکانی، همبستگی متقابل بین سه رنگ وجود دارد در شکل ۳ دیاگرام پراکنش سه رنگ با یکدیگر ارائه شده است. همانطورکه ملاحظه می‌شود بین رنگ‌ها با یکدیگر به‌صورت مستقیم ارتباطی خطی وجود دارد. ازاین‌رو می‌توان برای پیش‌گویی فضایی برای هرکدام از سه رنگ در هر پیکسل از مقادیر دیگر رنگ‌ها در آن پیکسل نیز استفاده کرد.

[۴]Primary
[۵]Secondary
[۶]Additive

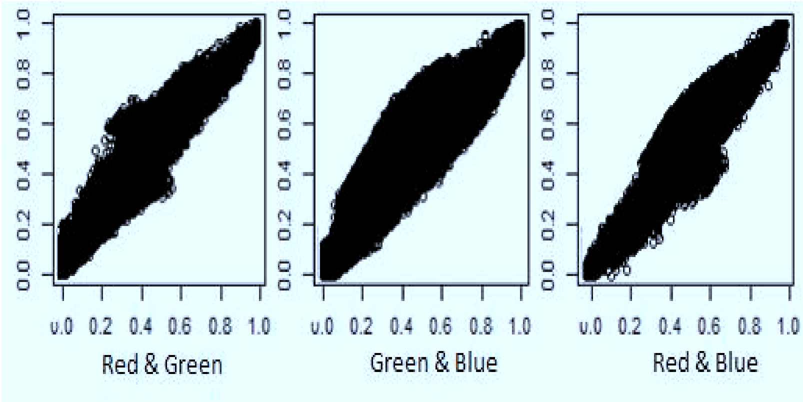
شکل ۲: نمایش تصویر برج آزادی با تفکیک RGB

وجود همبستگی فضایی در این سه تفکیک عکس کاملاً مشهود است اما جدای از همبستگی فضایی این سه رنگ ازلحاظ مکانی، همبستگی متقابل بین سه رنگ وجود دارد در شکل ۳ دیاگرام پراکنش سه رنگ با یکدیگر ارائه شده است. همانطورکه ملاحظه می‌شود بین رنگ‌ها با یکدیگر به‌صورت مستقیم ارتباطی خطی وجود دارد. ازاین‌رو می‌توان برای پیش‌گویی فضایی برای هرکدام از سه رنگ در هر پیکسل از مقادیر دیگر رنگ‌ها در آن پیکسل نیز استفاده کرد.

چنانچه رنگ‌هایی که در عکس برج آزادی بکار رفته است را به‌عنوان یک چارچوب فضایی با سه متغیر در نظر بگیریم، با استفاده از سه روش عنوان‌شده در این مقاله، به ترتیب نمونه‌های با حجم‌های ۲۰- ۴۰ -۷۰ – ۱۰۰ هزار نمونه را از این فضا گرفته‌ایم. در مرحله بعدی بافیت کردن یک توزیع فضایی به داده‌های نمونه‌گیری شده مقادیر سه رنگ را در نقاط دیگر فضا، پیش‌بینی کرده‌ایم. جدول زیر نتایج معیارهای ریشه توان دوم خطای پیش‌بینی فضایی در سه روش و با اندازه‌های نمونه‌ای متفاوت نشان داده‌شده است.

جدول ۱: مقادیر RMSE برای سه روش نمونه‌گیری فضایی و اندازه‌های مختلف نمونه

روش نمونه گیری	N=۲۰۰۰۰	N=۴۰۰۰۰	N=۷۰۰۰۰	N=۱۰۰۰۰۰
SRS	۰/۵۷۳	۰/۴۲۶	۰/۳۳۴	۰/۲۶۵
cubc	۰/۴۵۳	۰/۳۵۴	۰/۲۸۷	۰/۲۴۵
SBS	۰/۴۳۵	۰/۳۴۸	۰/۲۷۱	۰/۲۳۳



شکل ۳: نمودار پراکنش سه رنگ نسبت به یکدیگر

در جدول ۱ مقادیر RMSE برای سه روش نمونه‌گیری تصادفی ساده، نمونه‌گیری متعادل و نمونه‌گیری دومرحله‌ای نشان داده شده است. با توجه به‌اندازه‌های متفاوت نمونه و با توجه به روش‌های موجود می‌توان این‌گونه برداشت کرد که روش دومرحله‌ای توانسته خطای کمتری نسبت به روش‌های دیگر ایجاد کند.

۴ بحث و نتیجه‌گیری

روش نمونه‌گیری متعادل فضایی دومرحله‌ای معرفی و با استفاده از روش‌های پیش‌بینی در آمار فضایی با در نظر گرفتن نمونه‌ها توانستیم حجم تصویر را کم کنیم و بر اساس میانگین توان دوم خطای پیش‌گویی نشان داده شد روش نمونه‌گیری فضایی دومرحله‌ای توانسته بهتر از روش‌های مرسوم نمونه‌های بهینه را برای پیش‌گویی میدان تصادفی، انتخاب کند. دلیل این موضوع را نیز می‌توان به پخش کامل نمونه در فضای مطالعه و همچنین در نظر گفتن احتمال شمول بیشتر برای واحدهای حاشیه‌ای دانست. از طرفی با توجه به وابستگی‌های موجود بین سه رنگ قرمز، آبی و سبز در تصاویر RGB می‌توان بجای کریگینگ از روش کوکریگینگ استفاده کرد، که می‌تواند نتایج بهتری را حاصل نماید.

مراجع

محمدزاده م. (۱۳۹۱). آمار فضایی و کاربردهای آن، چاپ اول، مرکز نشر آثار علمی، دانشگاه تربیت مدرس، تهران.

Ansley, C. F. and Kohn, R. (1983), Exact Likelihood of Vector Autoregressive Moving Average Process with Missing or Aggregated Data, *Biometrika*, **70**, 275-278.

Chauvet, G. and Tille, Y. (2006), *A Fast Algorithm of Balanced Sampling*, To appear in Journal of Computational Statistique, INSEE.

Cressie, N. (1993), *Statistics for Spatial Data*, Revised edition, John Wiley, New York.

Cochran, W. G. and Waston, D. J. (1936), Empire J. Exp. *Agric*, **4**, 69-76.

Dessard, H. Bar-Hen, A. (2005), Experimental Design for Spatial Sampling Applied to the Study of Tropical Forest Regeneration, *Canadian Journal of Forest Research- Revue Canadienne De RechercheForestiere*, **35**, 1149-1155.

Dobbie,M. and Henderson, B. and Stevens, D. (2008), Sparse Sampling: Spatial Design for Monitoring Stream Networks, *Statistics*, **45**, 113-153.

Hedayat, A.S. and Majumdar, D. (1995), Generating Desirable Sampling Plans by the Technique of Trade-off in Experimental Design, *Journal of Statistical Planning and Inference*, **44**, 237-247.

McBratney, A. B. and Webster, R. (1981), Detection of Ride and Furrow Pattern by Spectral Analysis of Crop Yield, *International Statistic Review*, **49**, 45-52.

Stevens, Jr., D.L., and A.R. Olsen, (2004), Spatially-balanced Sampling of Natural Resources, *Journal of the American Statistical Association*, **99**, 262-277.

Royall, R. M. (1976), Likelihood Functions in Finite Population Sampling Theory, *Biometrika*, **63**, 605-614.

Royall, R. M. (1994), Discussion of ”Sample Surveys 1975- 1990; An Age of Reconciliation?” by T.M. F. Smith. *International Statistical Review*, **62**, 19- 21.

Salehi, M. (2004), Optimal sampling design under a spatial correlation model, *Journal of Statistical Planning and Inference*, **118**, 9-18.

Sarndal, C. E. and Swensson, B. and Wretman, J.(1992), *Model Assisted Survey Sampling*, Springer, New York.

Tille, Y. and Matei, A. (2005), *The R Package Sampling*, The Comprehensive R Archive Network, http: //cran. R- project. org/, Manual of the Contributed Packages.

Yates, F. (1934- 1935), Ann. *Eugenics*, **6**, 202- 213.