دانشکده علوم ریاضی
گروه آمار

پایان‌نامه

برای دریافت درجه کارشناسی ارشد در رشته
آمار، گرایش آمار ریاضی

عنوان

تشخیص داده‌های پرت فضایی

اساتید راهنما

دکتر داود شاهسونی - دکتر محمد کنشلو

استاد مشاور

دکتر حسین باغیشنی

دانشجو

مریم مقدم

بهمن ۱۳۹۲

به نام آن که جان را فکرت آموخت

پروردگارا...

نه می‌توانم مویشان را که در راه عزّت من سفید شد، سیاه کنم و نه برای دست‌های پینه بسته‌شان که ثمره تلاش برای افتخار من است، مرهمی دارم. پس توفیقم ده که هر لحظه شکرگزارشان باشم و ثانیه‌های عمرم را در عصای دست بودنشان بگذارم.

اگر تنهاترین تنها شوم، باز خدا هست

او جانشین همه نداشتن‌هاست...

تقدیم به خانواده عزیزم

سپاس‌گزاری

سپاس خدای را که سخنوران، در ستودن او بمانند و شمارندگان، شمردن نعمت‌های او ندانند و کوشندگان، حق او را گزاردن نتوانند، و سلام و دورد بر محمّد و خاندان پاك او، طاهران معصوم، هم آنان که وجودمان وامدار وجودشان است. به نشانه سپاس از الطافش، دستان مهربان‌ترین بندگانش، پدر و مادر عزیزم را بوسه می‌زنم.

اینک که این مرحله را پشت سر می‌گذارم، از تلاش‌ها، زحمات و راهنمایی‌های ظریف و ارزشمند اساتید فرزانه و گرانمایه‌ام، جنابان دکتر داود شاهسونی، دکتر محمد کنشلو و دکترحسین باغیشنی کمال تشکر را دارم. باشد که این خردترین، بخشی از زحمات آنان را سپاس گوید.

در پایان از جناب دکتر مهدی حسین‌پوری و سرکار خانم زهرا قیومی که دراین مدت کمک‌های زیادی به من کردند تشکر می‌کنم.

مریم مقدم

بهمن ۱۳۹۲

تعهدنامه

اینجانب **مریم مقدم** دانشجوی کارشناسی ارشد رشته **آمار** **دانشکده علوم ریاضی** دانشگاه شاهرود، نویسنده پایان‌نامه با عنوان **تشخیص داده‌های پرت فضایی** ، تحت راهنمایی **دکتر داود شاهسونی** و **دکتر محمد کنشلو** متعهد می‌شوم:

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان‌نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ‌جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه شاهرود متعلق دارد، و مقالات مستخرج با نام " دانشگاه شاهرود " یا " Shahrood University " به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده) شده است، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

مریم مقدم
بهمن ۱۳۹۲

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه شاهرود می‌باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان‌نامه بدون ذکر منبع مجاز نمی‌باشد.

چکیده

در ذخایر معدنی، تشخیص مقادیری که در دم راست توزیع واقع شده‌اند، از نظر اقتصادی اهمیت ویژه‌ای دارند؛ در نتیجه تشخیص مرز این‌گونه مشاهدات در صنعت معدن اهمیت بسزایی در اکتشاف و استخراج دارد. روش‌های موجود برای تشخیص داده پرت فضایی، مقدار این داده‌ها را ارائه می‌کنند و معیاری جهت شناسایی مرز مقادیر بالا ارائه نمی‌دهند. از این‌رو در این پایان‌نامه برای شناسایی مرز مقادیر بالا، روش مرز-بالا پیشنهاد می‌شود، که بر مبنای عدم وجود اثر مرزی استوار است. در صورت عدم وجود اثر مرزی، مقادیر بالای مرز، هیچ‌گونه تاثیر پذیری از مقادیر کمتر از مرز ندارد. در این روش، نسبت تغییرنگارهای شاخص متقابل به تغییرنگار شاخص برای مرزهای پیشنهادی محاسبه شده و درصورتی که تابعی از فاصله موقعیت‌ها نباشد، عدم وجود اثر مرزی را نشان می‌دهد. همچنین اگر تغییرنگار باقی‌مانده شاخص مرز پیشنهادی، اثر قطعه‌ای تام را نشان دهد، آنگاه مرز پیشنهادی به عنوان مرز مقادیر بالا انتخاب می‌گردد. ارزیابی عملکرد این روش و نیز برخی روش‌های تشخیص داده پرت، در یک مطالعه شبیه سازی انجام شده است. همچنین به عنوان یک مورد مطالعاتی واقعی، از این روش برای تشخیص مرز مقادیر بالای عیار داده‌های طلای ساری گونای در استان کردستان، استفاده شده و در نتیجه مرز ppm ۴/۶۳، به‌عنوان مرز عیار بالای طلای ساری گونای انتخاب گردید.

کلمات کلیدی: داده پرت فضایی، مقادیر بالا، اثر مرزی، تغییرنگار شاخص، تغییرنگار شاخص متقابل، روش مرز-بالا

مقاله مستخرج از پایان‌نامه

۱. مقدم، م.، کنشلو، م. و شاهسونی، د. (۱۳۹۲)، "کاربرد روش $Top-Cut$ در تعیین مرز مقادیر بالای عیار طلا "، اولین کنفرانس ملی مهندسی اکتشاف منابع زیرزمینی، ص ۲۱، شاهرود.

پیش‌گفتار

داده‌های فضایی مشاهداتی هستند که بر حسب موقعیت قرار گرفتن‌شان در فضای مورد مطالعه به یکدیگر وابسته‌اند. برآورد تغییرنگار و پیش‌گویی فضایی از مباحث مهم در آمار فضایی هستند. وجود داده‌های پرت فضایی یکی از مهم‌ترین عوامل بروز انحرافات، در این مباحث می‌باشد. بنابراین روش‌های تشخیص داده پرت فضایی، حائز اهمیت هستند. در این پایان نامه به شرح برخی از روش‌های تشخیص داده پرت فضایی می‌پردازیم. در ذخایر معدنی، منظور از داده پرت، مشاهداتی هستند که دارای مقادیر بالایی نسبت به سایر داده‌ها می‌باشند. مقادیر بالا (عیار بالا) در ذخایر معدنی اهمیت بسزایی در استخراج دارند. این‌گونه مقادیر در داده‌های طلا به صورت مرز شناخته می‌شوند. روش‌های موجود برای تشخیص داده‌های پرت فضایی، معیاری جهت تشخیص مرز مقادیر بالا ارائه نمی‌دهند. لذا در این پایان نامه به منظور تشخیص مرز مقادیر بالا، به شرح روشی به نام مرز–بالا می‌پردازیم که برگرفته از مدل مرز–بالا است.

مطالب فوق را در ۵ فصل بیان می‌کنیم:

- در فصل اول، پس از بیان مفاهیم و تعاریف آمار فضایی، به منظور شناخت اولیه زمین شناسی، به بیان تعاریف و مفاهیم آن پرداخته‌ایم و در آخر، اطلاعاتی در مورد طلا از جمله عیار، انواع کانسارهای طلا و ... ارائه نموده‌ایم.
- در فصل دوم، ابتدا به معرفی انواع داده پرت فضایی پرداخته و برخی روش‌های تشخیص آن‌ها را بیان نموده‌ایم. سپس مسایل مربوط به داده پرت در زمین آمار را بیان نموده و با مثال کوتاهی به اهمیت تشخیص مرز مقادیر بالا، در ذخایر معدنی پرداخته‌ایم.
- در فصل سوم، برای تشخیص مرز مقادیر بالا، روش مرز–بالا معرفی شده است. از آن‌جایی که این روش بر مبنای عدم وجود اثر مرزی استوار است قبل از معرفی آن، مفاهیم اثر مرزی و عدم وجود اثر مرزی را بیان نموده‌ایم.
- در فصل چهارم، در یک مطالعه شبیه سازی به ارزیابی عملکرد روش‌های تشخیص داده پرت فضایی و تشخیص مرز مقادیر بالا پرداخته‌ایم و همچنین به منظور تشخیص مرز عیار بالای یکی از معادن طلای ایران (ساری گونای)، از روش مرز–بالا استفاده نموده‌ایم. بدین منظور در ابتدا به تحلیل اکتشافی داده‌های معدن ساری گونای پرداخته‌ایم.

فهرست مطالب

فهرست تصاویر ز

۱ آشنایی با مفاهیم مقدماتی آمار فضایی ۱

۱.۱ مقدمه ۱

۲.۱ تعاریف و مفاهیم آمار فضایی ۲

۱.۲.۱ مانایی ۶

۲.۲.۱ تغییرنگار ۷

۳.۲.۱ هم‌تغییرنگار ۸

۴.۲.۱ برآورد تغییرنگار ۹

۵.۲.۱ تغییرنگار متقابل ۱۰

۶.۲.۱ همبستگی‌نگار ۱۱

۳.۱ تعاریف و مفاهیم زمین‌شناسی ۱۱

۴.۱ مقدمه‌ای بر مفاهیم مرتبط با عنصر طلا ۱۳

۱.۴.۱ مقدمه ۱۳

۲.۴.۱ تاریخچه ۱۳

۳.۴.۱ عیار طلا ۱۳

۴.۴.۱ ذخایر جهانی طلا ۱۴

۵.۱ مدل‌سازی ذخایر طلا ۱۵

۱.۵.۱ انواع کانسارهای طلا ۱۵

۲.۵.۱ تولید طلا ۱۸

۶.۱ نتایج ۱۸

۲ روش‌های تشخیص داده‌های پرت ۱۹

۱.۲ داده پرت فضایی ۱۹

۲.۲ تقسیم‌بندی روش‌های تشخیص داده‌های پرت ۲۱

۳.۲ تشخیص داده پرت فضایی با یک صفت کیفی ساده ۲۳
۱.۳.۲ نمودار پراکنش موران ۲۳
۲.۳.۲ نمودار پراکنش ۲۳
۳.۳.۲ الگوریتم Z ۲۵
۴.۳.۲ الگوریتم میانه ۲۵
۵.۳.۲ روش میانه-میانگین ۲۶
۴.۲ ابر تغییرنگار ۲۶
۵.۲ نمودار جعبه‌ای ۲۸
۶.۲ داده پرت در زمین‌آمار ۲۹
۱.۶.۲ برش زدن داده پرت ۳۰
۷.۲ نتیجه‌گیری ۳۳

۳ روش مرز-بالا برای تشخیص مرز مقادیر بالا **۳۵**
۱.۳ مقدمه ۳۵
۲.۳ اثر مرزی ۳۶
۳.۳ عدم وجود اثر مرزی ۳۶
۴.۳ روش مرز-بالا ۳۸
۱.۴.۳ مقدار بریده‌شده ۳۸
۲.۴.۳ مدل مرز-بالا ۳۹
۳.۴.۳ فرضیه‌های مدل مرز-بالا ۳۹
۴.۴.۳ انتخاب مقدار مرز-بالا ۴۲

۴ شبیه‌سازی و تحلیل اکتشافی داده‌های پروژه زرکوه **۴۵**
۱.۴ مطالعه شبیه‌سازی ۴۵
۲.۴ ارزیابی روش‌های تشخیص مرز مقادیر بالا ۴۷
۳.۴ ارزیابی عملکرد روش‌های تشخیص مقدار داده پرت فضایی ۴۸
۴.۴ تحلیل اکتشافی داده‌های زرکوه ۵۰
۵.۴ ساری گونای ۵۰
۶.۴ تحلیل اکتشافی داده مورد مطالعه ۵۰
۱.۶.۴ بررسی وجود همبستگی فضایی در داده‌ها ۵۱
۷.۴ نمودار جعبه‌ای ۵۴
۸.۴ ابر تغییرنگار ۵۴
۹.۴ روش مرز-بالا ۵۵

آ ۵۹

آ.۱ ساری گونای ۵۹

آ.۱.۱ زمین شناسی ساری گونای ۶۱

آ.۱.۲ نحوه کانه‌زایی در کانسار ۶۱

آ.۱.۳ زمین شناسی ساختاری ساری گونای ۶۲

ب کدهای مربوط به فصل‌های چهارم ۶۳

ب.۱ شبیه سازی ۶۳

ب.۲ روش مرز-بالا ۶۷

ب.۳ تحلیل اکتشافی ۶۹

مراجع ۷۰

واژه‌نامه فارسی به انگلیسی ۷۴

واژه‌نامه انگلیسی به فارسی ۷۶

فهرست تصاویر

۱.۱ موقعیت داده خاکستر زغال سنگ(نمونه‌ای از داده زمین‌آماری) ۴
۲.۱ (حق‌شناس تجن گوکه، ۱۳۸۹) نقشه متوسط اجاره‌بهای پرداختی ماهانه در شهر تهران به تفکیک ۲۲ منطقه شهرداری ۵
۳.۱ نمونه‌ای از الگوی نقطه‌ای ۵
۴.۱ نمودار تغییرنگار و هم‌تغییرنگار (محمدزاده، ۱۳۹۱) ۹

۱.۲ تقسیم‌بندی روش‌های تشخیص داده پرت ۲۱
۲.۲ نمودار پراکنش موران برای یک مجموعه داده فضایی ۲۴
۳.۲ نمودار پراکنش برای یک مجموعه داده فضایی ۲۴
۴.۲ نمودار پراکنش مقادیر خاکستر زغال‌سنگ در موقعیت‌های فضایی ۲۷
۵.۲ نمودار ابر تغییرنگار برای داده خاکستر زغال‌سنگ ۲۸
۶.۲ نمودار جعبه‌ای ۲۹
۷.۲ نمودار جعبه‌ای برای داده خاکستر زغال سنگ ۲۹
۸.۲ موقعیت داده پرت در بلوک ۳۱

۱.۳ نمونه‌ای از اثر مرزی ۳۷
۲.۳ نسبت واریوگرام‌های شاخص برای مرزهای ۵، ۱۰ و ۲۰ . . . . . ۴۴
۳.۳ تغییرنگارهای باقی‌مانده‌های شاخص برای $5-10g/t$ و $10-20g/t$ ۴۴

۱.۴ شبکه منظم داده‌های شبیه‌سازی شده ۴۶
۲.۴ الف: بافت‌نگار نمونه‌های شبیه‌سازی شده ب: بافت‌نگار نمونه‌های شبیه‌سازی شده آلوده به داده‌های پرت ۴۶
۳.۴ معیار ارزیابی $TPrate$ برای روش‌های نمودار جعبه‌ای و مرز-بالا جهت تشخیص مرز مقادیر بالا ۴۷
۴.۴ نمودارهای معیارهای سنجش برای الگوریتم میانه در $k=2$ ۴۸
۵.۴ نمودارهای معیارهای سنجش برای الگوریتم میانه در $k=8$ ۴۸

۶.۴ نمودارهای معیارهای سنجش برای الگوریتم میانه در $\alpha/۲ = ۰/۰۰۱$ ۴۹
۷.۴ نمودارهای معیارهای سنجش برای الگوریتم میانه در $\alpha/۲ = ۰/۰۰۵$ ۴۹
۸.۴ نمودارهای معیارهای سنجش برای الگوریتم میانه در $\alpha/۲ = ۰/۰۲۵$ ۴۹
۹.۴ نمودارهای معیارهای سنجش برای الگوریتم میانه و روش میانه میانگین ۴۹
۱۰.۴ موقعیت کانسار طلای ساری گونای در ایران ۵۰
۱۱.۴ پراکندگی داده‌ها نسبت به محورهای مختصات ۵۲
۱۲.۴ الف: نمودار چندک چندک داده‌ها، ب: نمودار چندک چندک داده‌ها همراه با کاهش چولگی ۵۳
۱۳.۴ الف: نمودار جعبه‌ای داده‌های عیار طلا، ب:نموار جعبه‌ای داده‌های عیار طلا همراه با کاهش چولگی ۵۵
۱۴.۴ نمودارهای ابر تغییرنگار داده‌های عیار طلای زرکوه ۵۵
۱۵.۴ الف: نمودار $\frac{\gamma I_{(۲/۶,۴/۶۳)(h)}}{\gamma I_{(۲/۶)(h)}}$ ب: نمودار $\frac{\gamma I_{(۴/۶۳,۶/۵۹)(h)}}{\gamma I_{(۴/۶۳)(h)}}$ ج: نمودار $\frac{\gamma I_{(۶/۵۹,۸/۴۱۲)(h)}}{\gamma I_{(۶/۵۹)(h)}}$ . . ۵۷
۱۶.۴ تغیرنگارهای باقی‌مانده شاخص برای ۴/۶۳-۲/۶، ۶/۵۹-۴/۶۳ و ۸/۴۱۲-۶/۵۹ ۵۷

آ.۱ نمای منطقه پروژه ساری گونای از جنوب غربی ۶۱
آ.۲ زون‌های پرعیار و کم عیار ساری گونای ۶۲

فصل ۱

آشنایی با مفاهیم مقدماتی آمار فضایی

۱.۱ مقدمه

در روش‌های معمول آماری، اغلب فرض می‌شود که مشاهدات حاصل از نمونه استخراج شده از جامعه مستقل از یکدیگرند. اما در بسیاری از موارد، مشاهدات مستقل نیستند و برحسب موقعیت قرار گرفتن‌شان در فضای مورد مطالعه، به یکدیگر وابسته هستند. اگر این وابستگی تابعی از فاصله بین موقعیت‌های مشاهدات باشد، به‌گونه‌ای که مشاهدات نزدیک به هم، وابسته‌تر و مشاهداتی که دور از یکدیگر هستند، وابستگی کمتری داشته باشند، آنگاه این‌گونه مشاهدات را داده‌های فضایی[۱] می‌نامند (محمدزاده، ۱۳۹۱).

مبانی آمار فضایی و گسترش روش‌های آن، به دنبال پژوهش‌های پروفسور ماترون[۲] بنا نهاده شد. این پژوهشگر فرانسوی با انتشار مقاله بنیادین خود توانست شاخه جدیدی از علم آمار با عنوان "زمین‌آمار"[۳] را پایه‌گذاری کند. همچنین ماترون، روش کریگینگ[۴] را، که منتسب به دی.جی.کریگ[۵]، مهندس معدن آفریقای جنوبی است، مطرح کرد. دیوید[۶] (۱۹۷۷) نیز در مورد پیش بینی‌های فضایی ذخایر معدنی، مطالب بسیاری را ارائه کرده است. کرسی[۷] (۱۹۹۳) با انتشار کتاب "آمار برای داده‌های فضایی" در تکامل این شاخه از علوم قدم بزرگی را برداشت.

نخستین تجربه‌ها برای به‌کارگیری روش‌های آماری در پیش‌گویی ذخایر معادن، با شناسایی مقدماتی الگوی توزیع طلا در معادن آفریقای جنوبی آغاز شد. در آن زمان نکته جالب توجه محققان این بود

---

[۱] Spatial Data
[۲] Matheron
[۳] Geostatistics
[۴] Kriging
[۵] Danniel Gerhardus Krige
[۶] David
[۷] Cressie

که در یک منطقه معدنی، بخش‌های پرعیار و کم‌عیار به‌طور معمول در کنار یکدیگر قرار دارند و باید نوعی رابطه میان این بخش‌ها وجود داشته باشد (محمدزاده، ۱۳۹۱). بخش‌های پر عیار (مقادیر بالا) در ذخایر معدنی، داده پرت[۱] نامیده می‌شود. در راستای تشخیص داده پرت فضایی، تحقیقات زیادی انجام شده است. از جمله کو[۲] و همکاران (۲۰۰۷)، الگوریتمی را بر مبنای k-نزدیکترین همسایگی، برای تشخیص نقاط یا ناحیه‌های معمولی که خیلی شبیه به داده‌هایی هستند که در همسایگی‌شان قرار دارند، در نظر گرفتند. الگوریتم پیشنهادی آن‌ها، به علت دقت در تشخیص داده‌های پرت فضایی و توانایی تشخیص داده‌های پرت ناحیه‌ای، نسبت به روش‌های دیگر تشخیص داده پرت فضایی برتری دارد. همچنین الگوریتم k نزدیکترین همسایگی مبنا[۳] ($BKNN$)، توسط کاو[۴] (۲۰۱۳a) و همکاران که بر مبنای اصول نمودار KNN است، به منظور رفع این مشکل که مشاهدات فضایی ممکن است توسط همسایگی‌های غیرمعمول‌شان، تاثیرپذیر باشند ارائه گردید. کاو و همکاران ،(۲۰۱۳b) نیز نادیده گرفتن تاثیر مشخصه‌های فضایی بر روی درجه پرت بودن داده‌های پرت فضایی را مشکل روش‌های رایج تشخیص داده پرت فضایی می‌داند و برای این منظور ایده‌ای بر مبنای داده پرت وزن‌دار بیان می‌کند. این مطالعه، تاثیرات مشخصه‌های فضایی را بر روی محاسبه درجه پرتی فضایی تشخیص می‌دهد و این تاثیرات را با مشخصه‌های غیرفضایی ترکیب می‌کند و دو الگوریتم z-مقدار توسعه یافته[۵] و الگوریتم تفاضل وزن‌دار[۶] را پیشنهاد می‌کند.

از آنجاکه بخش‌های پر عیار و کم عیار معدن در کنار یکدیگر قرار دارند، لذا تمرکز ما در این پایان‌نامه، مطالعه روش تشخیص مرز مقادیر بالا برای ذخایر معدنی است که از این روش، برای تشخیص مرز بالای عیار طلا در یکی از معادن طلای ایران استفاده شده است.

۲.۱ تعاریف و مفاهیم آمار فضایی

برای مدل‌سازی و تجزیه و تحلیل داده‌های فضایی، نیاز به تعریف میدان تصادفی است.[۷]

تعریف ۱.۲.۱. میدان تصادفی: فرض کنید $\mathrm{s} \in \mathrm{R}^d$، موقعیت داده در فضای اقلیدسی dبعدی ($\mathrm{d} \geq ۱$) و $\mathrm{Z(s)}$ یک کمیت تصادفی در موقعیت s باشد. حال اگر s روی مجموعه اندیس‌گذار $\mathrm{D} \subset \mathrm{R}^d$ تغییرکند، موجب تولید میدان تصادفی

$$\{\mathrm{Z(s)}; \mathrm{s} \in \mathrm{D}\}, \qquad (۱.۱)$$

[۱] Outlier
[۲] Kou
[۳] Based K-Nearest Neighbor
[۴] Cao
[۵] Improved Z-value
[۶] Weighted difference algorithm
[۷] مفاهیم این بخش، از کتاب دکتر محمدزاده (۱۳۹۱) اقتباس گردیده است.

می‌شود که $\{Z(s); s \in D\}$ یک تحقق بیان شده از میدان تصادفی (۱.۱) است. توابع میانگین، کوواریانس و واریانس میدان تصادفی $Z(\cdot)$ به‌ترتیب به‌صورت زیر بیان می‌شوند:

$$\begin{aligned}
\mu(s) &= E[Z(s)], \quad s \in D,\\
C(s_i, s_j) &= Cov[Z(s_i), Z(s_j)]\\
&= E[(Z(s_i) - \mu(s_i))(Z(s_j) - \mu(s_j))], \quad s_i, s_j \in D,\\
Var[Z(s)] &= E[Z(s) - \mu(s)]^2 = C(s, s).
\end{aligned}$$

هر میدان تصادفی (۱.۱) را می‌توان به‌صورت

$$Z(s) = \mu(s) + \delta(s), \quad s \in D,$$

تجزیه کرد که در آن $\mu(\cdot)$، تغییرات بزرگ‌مقیاس[۱] یا روند[۲] میدان تصادفی است، که ممکن است ناشی از تغییرپذیری **بین** موقعیت‌های مشاهده شده باشد، و $\delta(\cdot)$، فرآیند خطا یا تغییرات کوچک‌مقیاس[۳] میدان تصادفی است، که ممکن است ناشی از خطای اندازه‌گیری یا تغییرپذیری در **درون** موقعیت مشاهده‌شده باشد.

مثال ۲.۲.۱. فرض کنید برای یک میدان تصادفی در پنج موقعیت

$$D = \{s_1 = (1, 1), s_2 = (2, 2), s_3 = (3, 4), s_4 = (1, 4), s_5 = (3, 1)\}$$

مقادیر زیر مشاهده شده‌اند.

$$z(s_1) = 1, z(s_2) = 2, z(s_3) = 20, z(s_4) = 4, z(s_5) = 3$$

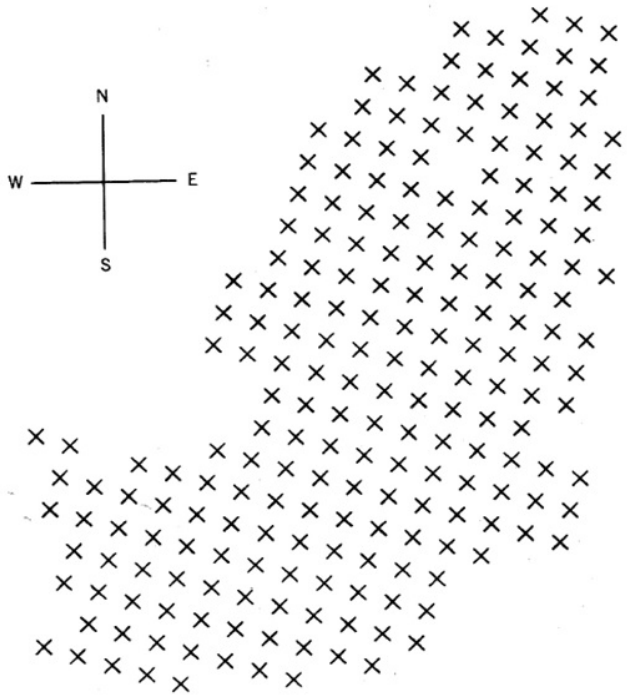
در این‌صورت $\{z(s_1), z(s_2), z(s_3), z(s_4), z(s_5)\}$، تحققی از میدان تصادفی $\{Z(s); s \in D\}$ است.

در آمار فضایی، متغیرهای مورد اندازه‌گیری ممکن است گسته یا پیوسته باشند. فضای موقعیت یا مکان مشاهدات نیز ممکن است پیوسته یا گسسته، نقطه‌ای یا ناحیه‌ای، منظم یا نامنظم باشد. با توجه به انواع موقعیت‌ها، مشاهدات فضایی به سه گروه عمده، داده‌های زمین‌آماری[۴]، داده‌های مشبکه‌ای[۵] و الگوهای نقطه‌ای[۶] تقسیم می‌شوند (کرسی، ۱۹۹۳).

داده‌های زمین‌آماری: این نوع داده‌ها در موقعیت‌های ثابت و مشخص، در ناحیه‌ای پیوسته مشاهده می‌شوند. مکان و متغیر مورد مطالعه، ممکن است گسسته یا پیوسته باشد. به‌عنوان مثال، برای حالت پیوسته می‌توان غلظت مواد معدنی در داخل یک کانسار[۷]، مقدار ریزش باران ثبت‌شده در ایستگاه‌های

[۱] Large scale variation
[۲] Trend
[۳] Small scale variation
[۴] Geostatistical data
[۵] Lattice data
[۶] Point pattern
[۷] Mineral deposit

هواشناسی، و برای حالت گسسته می‌توان تعداد نوعی جانور دریایی در مکان‌های نمونه‌گیری‌شده در طول یک ساحل را نام برد. داده‌های مربوط به موقعیت خاکستر زغال سنگ[۱] نمونه‌ای از داده‌های زمین‌آماری، در شکل (۱.۱) نمایش داده شده است.



شکل ۱.۱: موقعیت داده خاکستر زغال سنگ(نمونه‌ای از داده زمین‌آماری)

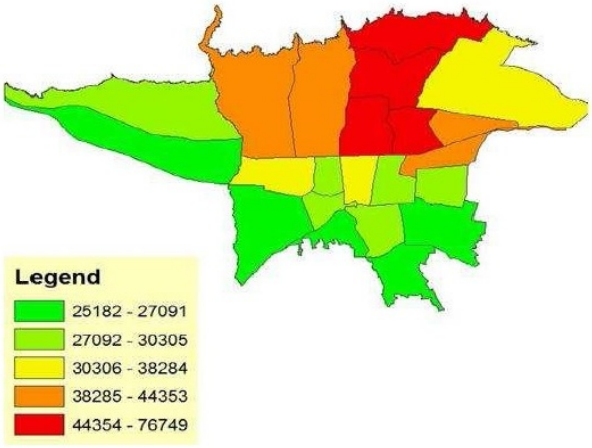
داده‌های مشبکه‌ای: این نوع داده‌ها مربوط به مکان‌های ناحیه‌ای هستند، که این مکان‌ها ممکن است منظم یا نامنظم باشند. تصاویر ماهواره‌ای از سطح زمین که در آن‌ها سطح زمین به تعدادی عنصر تصویر[۲] کوچک به‌صورت مشبکه منظم در R^2 تقسیم شده، مثالی برای داده‌های مشبکه‌ای **منظم** است. تعداد افراد مبتلا به سرطان در مناطق مختلف خدمات درمانی کشور، مثالی برای داده‌های مشبکه‌ای **نامنظم** است. در این حالت در هر شهرستان، کل افراد مبتلا به سرطان، در مرکز خدمات درمانی آن شهرستان، در نظر گرفته می‌شوند. به‌طور معمول، هدف از تحلیل داده‌های فضایی مشبکه‌ای، مدل‌بندی احتمالاتی مشاهدات است، در صورتی که در زمین‌آمار، پیش‌گویی مقدار متغیر در یک موقعیت جدید موردنظر است. برای مثال نقشه متوسط اجاره‌بهای پرداختی ماهانه در شهر تهران به تفکیک ۲۲ منطقه شهرداری، در شکل (۲.۱) نمایش داده شده است.

الگوی نقطه‌ای: در این گروه از داده‌ها، مکان یا موقعیت مشاهده‌شده، متغیر تصادفی موردنظر است. الگوی نقطه‌ای، شامل تعدادی متناهی از مکان‌ها در یک ناحیه است که در آن‌ها یک صفت خاص اندازه‌گیری می‌شود. به‌طور معمول، الگوی‌های نقطه‌ای به سه دسته به‌طور کامل تصادفی فضایی[۳]

[۱] Coal Ash Data

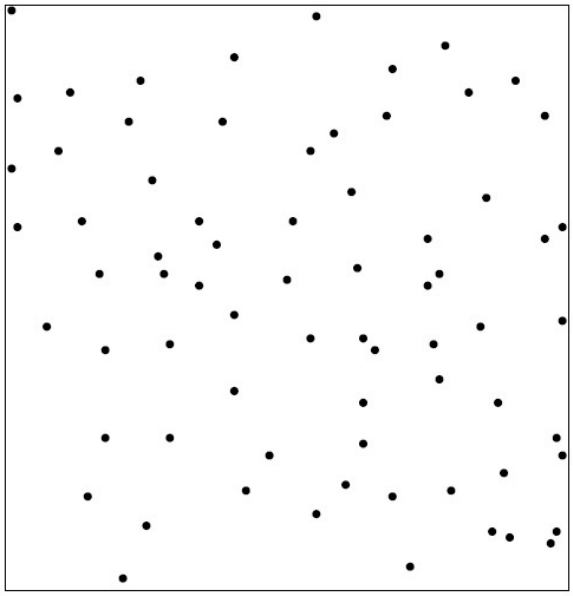
[۲] Pixel

[۳] Complete spatial randomness



شکل ۲.۱: (حق‌شناس تجن گوکه، ۱۳۸۹) نقشه متوسط اجاره‌بهای پرداختی ماهانه در شهر تهران به تفکیک ۲۲ منطقه شهرداری

(CSR)، منظم[۱] و خوشه‌ای[۲] تقسیم گردیده و به مدل‌بندی آن‌ها اقدام می‌شود. یک مثال برای این نوع مشاهدات، موقعیت گونه‌ای از درختان در یک ناحیه جنگلی یا موقعیت مراکز زلزله است. نمونه‌ای از الگوی نقطه‌ای که موقعیت آن‌ها به‌صورت تصادفی قرار گرفته‌اند، در شکل (۳.۱) نمایش داده شده است.



شکل ۳.۱: نمونه‌ای از الگوی نقطه‌ای

در آمار کلاسیک برای نشان دادن میزان پراکندگی و وابستگی خطی داده‌ها، به ترتیب از واریانس

[۱] Regularity
[۲] Clustering

و کوواریانس استفاده می‌شود. در آمار فضایی بر حسب این‌که مشاهدات فضایی، مربوط به کدامیک از دسته‌های فوق باشند، بررسی و تحلیل آن‌ها نیز متفاوت است. به‌عنوان مثال، محاسبه وابستگی فضایی برای داده‌هایی که از نوع زمین‌آماری باشند، بر حسب تغییرنگار[۱] و برای داده‌ها مشبکه‌ای، بر اساس ماتریس همسایگی محاسبه می‌شود.

۱.۲.۱ مانایی

تعریف ۳.۲.۱. مانایی قوی [۲]: $Z(s)$ مانا است، اگر برای فاصله h، و همه موقعیت‌های $s_1, \cdots, s_n$ که n متناهی است، توزیع توام $Z(s_1), \cdots, Z(s_n)$ همانند توزیع توام $Z(s_1 + h), \cdots, Z(s_n + h)$ باشد. یعنی به بیان دیگر، در صورت انتقال موقعیت‌های $s_1, \cdots, s_n$ در راستای $h \in R$ توزیع توام $Z(s_1), \cdots, Z(s_n)$ تغییر نمی‌کند (کاکس و میلر[۳]، ۱۹۶۵).

تعریف ۴.۲.۱. مانایی مرتبه دوم[۴]: به دو دلیل، یکی مشکل بودن آزمودن مانایی قوی و دیگری این‌که در این مانایی (مانایی قوی) وجود گشتاورها الزامی نیست، اغلب نوع ضعیف‌تری از آن که با عنوان مانایی مرتبه دوم شناخته شده است، به‌کار گرفته می‌شود.

$Z(s)$ مانای مرتبه دوم است، اگر$\mathrm{Cov}(Z(s+h), Z(s))$ وجود داشته باشد و تنها به h بستگی داشته باشد یعنی میزان تغییرپذیری یا همبستگی متغیر در تاخیر h، تابعی از فاصله بین موقعیت‌های آن‌ها می‌باشد. این شرط، وجود $\mathrm{Var}(Z(s))$ و عدم وابستگی آن به S را الزامی می‌دارد یعنی واریانس میدان تصادفی به موقعیت فضایی بستگی ندارد و تغییرپذیری میدان همه‌جا یکسان است. به همین ترتیب $E(Z(s))$ نیز وجود دارد و به S وابسته نمی‌باشد (کاکس و میلر، ۱۹۶۵). بدیهی است، مانایی به معنای برقراری مانایی مرتبه دوم نمی‌باشد بدین معنا که یک تابع تصادفی مانا ممکن است به علت عدم وجود گشتاورهای مراتب اول دوم، مانای مرتبه دوم نباشد (میرس[۵]، ۱۹۸۹).

تعریف ۵.۲.۱. میدان تصادفی $Z(.)$ مانای ذاتی[۶] نامیده می‌شود، هرگاه

۱. میانگین میدان تصادفی ثابت یا مستقل از s باشد، یعنی $E[Z(s)] = \mu$

۲. واریانس عبارت $(Z(s) - Z(s+h))$ فقط تابعی از فاصله موقعیت‌های s و $s+h$ باشد، یعنی

$$\mathrm{Var}[Z(s+h) - Z(s)] = 2\gamma(h).$$

(محمدزاده، ۱۳۹۱).

---

[۱] Variogram

[۲] Strong stationary

[۳] Cox and Miller

[۴] Second order stationary

[۵] Myers

[۶] Intrinsic stationary