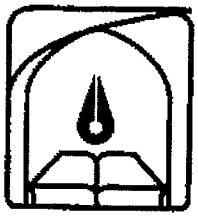
۱۳۸۱ /۱۲/ ۱۰



دانشگاه تربیت مدرس

دانشکده علوم پزشکی

رسالهٔ دورهٔ دکترای تخصصی (Ph.D.) در رشته آمار زیستی

عنوان

تجزیه و تحلیل داده‌های فضایی نامانا و کاربرد آن در همه‌گیری‌شناسی جغرافیایی بیماریها

نگارش

یداله واقعی

استاد راهنما

دکتر محسن محمدزاده

اساتید مشاور

دکتر انوشیروان کاظم‌نژاد

دکتر سقراط فقیه‌زاده

پاییز ۱۳۸۱

« تأییدیه اعضای هیأت داوران حاضر در جلسه دفاع از رساله دکتری

اعضای هیئت علمی داوران نسخه نهائی رساله خانم / آقای یداله واقعی **رشته** : آمارزیستی

تحت‌عنوان عنوان‌رساله: تجزیه و تحلیل داده‌های فضایی ناماناو کاربرد آن در همه‌گیری‌شناسی جغرافیایی بیماریها

را از نظر فرم و محتوی بررسی نموده و پذیرش آنرا برای تکمیل درجه دکتری پیشنهاد می‌کنند.

نام و نام خانوادگی و امضاء اعضای هیأت داوران:

آقای دکتر محسن محمدزاده - استاد راهنما

آقای دکتر انوشیروان کاظم نژاد - استاد مشاور

آقای دکتر سقراط فقیه‌زاده - استاد مشاور

آقای دکتر غلامرضا بابایی - استادناظرونماینده‌شورای تحصیلات تکمیلی

آقای دکتر محمدرضامشکانی - استادناظر

آقای دکتر محمودمحمودی‌فراهانی - استادناظر

آقای دکتر علیرضا حیدرنیا - استاد ناظر

بسمه‌تعالی

آیین‌نامهٔ چاپ پایان‌نامه (رساله)های دانشجویان دانشگاه تربیت مدرس

نظر به اینکه چاپ و انتشار پایان‌نامه (رساله)های تحصیلی دانشجویان تربیت مدرس، مبین بخشی از فعالیتهای علمی ـ پژوهشی دانشگاه است بنابراین به منظور آگاهی و رعایت حقوق دانشگاه، دانش آموختگان این دانشگاه نسبت به رعایت موارد ذیل متعهد می‌شوند:

ماده ۱ در صورت اقدام به چاپ پایان‌نامه (رساله)ی خود، مراتب را قبلاً به طور کتبی به «دفتر نشر آثار علمی» دانشگاه اطلاع دهد.

ماده ۲ در صفحه سوم کتاب (پس از برگ شناسنامه)، عبارت ذیل را چاپ کند:

«کتاب حاضر، حاصل رساله دکترای نگارنده در رشتهٔ آمار زیستی است که در سال ۱۳۸۱ در دانشکدهٔ علوم پزشکی دانشگاه تربیت مدرس به راهنمایی جناب آقای دکتر محسن محمدزاده و مشاورهٔ جناب آقای دکتر انوشیروان کاظم‌نژاد و مشاورهٔ جناب آقای دکتر سقراط فقیه‌زاده از آن دفاع شده است».

ماده ۳ به منظور جبران بخشی از هزینه‌های انتشارات دانشگاه، تعداد یک درصد شمارگان کتاب (در هر نوبت چاپ) را به «دفتر نشر آثار علمی» دانشگاه اهدا کند. دانشگاه می‌تواند مازاد نیاز خود را به نفع مرکز نشر در معرض فروش قرار دهد.

ماده ٤ در صورت عدم رعایت مادهٔ ۳، ۵۰٪ بهای شمارگان چاپ شده را به عنوان خسارت به دانشگاه تربیت مدرس، تأدیه کند.

ماده ۵ دانشجو تعهد و قبول می‌کند در صورت خودداری از پرداخت بهای خسارت، دانشگاه می‌تواند خسارت مذکور را از طریق مراجع قضایی مطالبه و وصول کند؛ به علاوه به دانشگاه حق می‌دهد به منظور استیفای حقوق خود، از طریق دادگاه، معادل وجه مذکور در مادهٔ ٤ را از محل توقیف کتابهای عرضه شدهٔ نگارنده برای فروش، تأمین نماید.

ماده ٦ اینجانب یدالله واقعی دانشجوی رشته آمار زیستی مقطع دکترا تعهد فوق و ضمانت اجرایی آن را قبول کرده، به آن ملتزم می‌شوم.

نام و نام خانوادگی : یدا... واقعی

تاریخ و امضا : ۸۱/٦/۱۳

تقدیم به :

پدر و مادر بزرگوار

همسر مهربان

و فرزند عزیزم، مریم

که در دوران تحصیل اینجانب سختی‌ها و زحمات زیادی را متقبل شده و مشوقِ من در طی راه بوده‌اند. بی‌شک همدلی، همراهی و دعای خیر این عزیزان، مایـــه دلگرمی و پیشرفت اینجانب بوده است.

قدردانی

من لم یشکر المخلوق لم یشکر الخالق

خداونـد منان را سپاس می‌گویم که به من توفیق توانایی طی این مرحله از تحصیل را عطا کرد. بر خود لازم می‌دانم از کلیه کسانی که در طی مراحل مختلف تحصیل به گونه‌ای من را یاری داده‌اند تقدیر و تشکر نمایم. از :

- استاد راهنمای گرانقدر و فرزانه جناب آقای دکتر محمدزاده که در طی بیش از سه سال هدایت علمی این رساله را بـر عهده داشته و با نکته‌سنجی، صرف وقت فراوان و ارائه ایده‌ها و رهنمودهای علمی خود بر ارزش و غنای این تحقیق افزودند؛
- اسـاتید مشـاور ارجمند آقایان دکتر کاظم‌نژاد و دکتر فقیه‌زاده که راه شروع و پیشبرد این تحقیق را هموار ساخته و از محضر درس آنها بهره برده‌ام؛
- اسـاتید عزیـز آقایان دکتر ارقامی، دکتر بابایی، دکتر پاشا، دکتر گرامی، دکتر محمدرضا مشکانی و دکتر محمودی که اساتیدم در این مقطع از تحصیل بوده‌اند؛
- اسـاتید معظم آقایان دکتر پارسیان، دکتر پاشا، دکتر طالبی، دکتر موسوی و دکتر نواب‌پور که در مقطع کارشناسی ارشد از محضرشان کسب علم نموده‌ام؛
- آقایـان دکـتر آذرنـوش، دکـتر بزرگ‌نیا، دکتر شاهکار، دکتر صادقی، دکتر طباطبایی، دکتر نیرومند و سایر اساتید گـروههای آمـار و ریاضی دانشگاه فردوسی مشهد که دردوره کارشناسی اصول و مبانی علم آمار را به خوبی به من آموختند؛
- داوران محـترم این رساله، آقایان دکتر مشکانی ، دکتر محمودی ، دکتر بابائی و دکتر حیدرنیا که با صرف فقط و دقت نظر راهنمایی‌های علمی و فنی را در جهت بهبود کمی و کیفی نگارش رساله ارائه نموده‌اند؛
- مسـؤولین و کارشناسـان محـترم اداره کـل مبارزه با بیماریها که داده‌های مورد استفاده در این رساله را در اختیار من قرار داده‌اند؛
- آقایـان دکـتر فقیهـی و دکـتر شـادرخ ایزدی و آقای بلادی موسوی که با توصیه‌ها و راهنمایی‌های علمی من را مفتخر گردانده‌اند؛
- آقایـان حسـن رنجـی، غلامرضـا ایزدی، محمد نجاری، سعید فتحی و خانم افسانه یزدانی که راهنمایی‌های فنی فراوانی در زمینه نصب و کاربرد نرم‌افزارهای مورد نیاز ارائه نموده‌اند؛
- همچـنین از آقای ریبه‌ایرو (Paulo Justiniano Ribeiro Jr) از دانشگاه لنکستر به خاطر راهنمایی‌های فنی و علمی ارزنده و پاسخگویی سریع به سؤالات و ابهامات مکرر در مورد بسته نرم‌افزاری GeoR ؛
- و خانم محبوبه کاظمی که زحمت تایپ این پایان‌نامه را قبول فرموده‌اند،

و از همـه کسـانی کـه به نحوی در پیشبرد این رساله مؤثر بوده‌اند قدردانی و سپاسگزاری فراوان نموده و از خداوند متعال برای آنها توفیق و سلامتی در جهت ارائه خدمت طلب می‌نمایم.

اجرهم عندا...

چکیده:

داده‌های فضایی داده‌های همبسته‌ای هستند که از موقعیت‌های مختلف در فضای d بعدی(معمولاً d=۲) جمع‌آوری می‌شوند. ساختار همبستگی این داده‌ها نوعاً به فاصله بین موقعیتهای آنها بستگی دارد و تحت شرط مانایی بوسیله تغییرنگار یا هم‌تغییرنگار برآورد و مدلسازی می‌شود. وجود روند فضایی در داده‌ها سبب نامانایی شده و برآورد تغییرنگار حاصل از این داده‌ها نااریب خواهد بود. چنانچه تغییرنگار داده‌ها در جهت‌های مختلف متفاوت باشد داده‌ها ناهمسانگرد نامیده می‌شود. ناهمسانگردی سبب پیچیدگی رفتار و مدلسازی تغییرنگار می‌شود. وجود داده‌های دورافتاده فضایی نیز برآورد تغییرنگار و پیشگویی‌های مبتنی بر آن را تحت تاثیر قرار می‌دهد. به همین دلیل در این رساله روشهای مختلفی برای مشکلات ناشی از نامانایی، ناهمسانگردی، و داده‌های دورافتاده در تجزیه و تحلیل داده‌های فضایی پرداخته شده و کاربرد آنها در پیشگویی فضایی میزان بروز بیماریها نشان داده شده است.

در این رساله نخست روشهای برآورد و مدلسازی تغییرنگار داده‌های مانا و همسانگرد معرفی شده است. سپس برآورد تغییرنگار بر اساس داده‌های روندزدوده به عنوان یک راه حل برای مشکل نامانایی ناشی از وجود روند در داده‌ها مورد بررسی قرارگرفته و همچنین استفاده از تابع کواریانس داده‌های روندزدوده به عنوان یک راه حل برتر که به نحوهٔ روندزدایی داده‌ها بستگی ندارد مطرح شده است. علاوه بر آن روشهای مختلف کریگیدن برای پیشگویی فضایی معرفی شده و به روشهای مختلف شناسایی و برخورد با داده‌های فضایی پرداخته شده و نتایج کاربرد این روشها روی داده‌های واقعی نشان داده شده است.

برای داده‌های ناهمسانگرد ممکن است یک یا چند پارامتر، یا در حالت کلی‌تر مدل تغییرنگار به جهت وابسته باشد. به منظور لحاظ کردن ناهمسانگردی در برآورد تغییرنگار مدلهای مختلفی برای ناهمسانگردی ناشی از تغییر یکی از پارامترهای بُرد یا ازاره ارائه شده است. همچنین یک مدل کلی‌تر پیشنهاد شده که در آن همهٔ پارامترهای مدل می‌توانند با جهت تغییر کنند. به کمک شبیه‌سازی دقت این مدل با مدل همسانگرد مقایسه شده و در پیشگویی به روش کریگیدن به کار رفته است.

به منظور نمایش کاربرد روشهای ارائه شده در همه‌گیری شناسی جغرافیایی بیماریها ساختار همبستگی داده‌های سل ریوی اسمیر مثبت سالهای ۱۳۷۷ و ۱۳۷۸ در ۲۶۲ شهرستان کشور، تحت دو فرض همسانگردی و ناهمسانگردی مدلسازی شده و میزان بروز این بیماری در ۲۰ شهرستان فاقد اطلاع با کریگیدن پیشگویی شده است. علاوه بر این فصل آخر به تهیه نقشه آماری بیماریها، که از مباحث مهم در همه‌گیری شناسی جغرافیایی بیماریهاست، پرداخته شده و به طور اخص نقشه بروز سل ریوی اسمیر مثبت به روش کریگیدن عام برای سالهای ۱۳۷۷ و ۱۳۷۸ ارائه شده و مورد مقایسه قرار گرفته‌اند و نقشه انحراف معیار کریگیدن نیز تهیه شده است.

از آنجا که داده‌های سل ریوی اسمیر مثبت به دلایلی از جمله عدم مراجعه بعضی از بیماران به شبکه بهداشتی یا تشخیص نادرست بیماری تحت تاثیر کم‌ثبتی و احیاناً خطاهای دیگری می‌باشند، که میزان آنها مشخص نیست، نتایج عددی پیشگویی و نقشه‌های حاصل از آن به طور غیر مستقیم تحت تاثیر این خطاها می‌باشد. لذا تأکید این رساله بیشتر بر فنون و روشهای ارائه شده می‌باشد و از داده‌های یاد شده صرفاً برای بیان و نمایش کاربرد روشها استفاده شده است. در عین حال فنون و روشهایی که در این رساله ارائه شده‌اند کلی بوده و برای تجزیه و تحلیل داده‌های فضایی در حیطه‌های زمین‌شناسی، هوا شناسی، بوم‌شناسی، همه‌گیری‌شناسی، بهداشت محیط، مهندسی محیط زیست، دورسنجی و کشاورزی قابل استفاده‌اند.

واژه‌های کلیدی: *داده‌های فضایی، نامانایی، ناهمسانگردی، داده‌های دورافتاده، کریگیدن، همه‌گیری شناسی جغرافیایی، نقشه بیماری*

فهرست مطالب

عنـــــــــوان مطلب . . . . . صفحه

فصل اول : مقدمه و کلیات

۱-۱ تاریخچه ۱
۱-۲ اهداف و محتویات رساله ٤
۱-۳ داده‌های فضایی ٥
۱-۳-۱ انواع داده‌های فضایی ٦
۱-٤ میدان تصادفی ۸
۱-٥ تعاریف ۹
۱-٥-۱ مانایی ۱۰
۱-٥-۲ تغییرنگار ۱۱
۱-٥-۳ هم‌تغییرنگار ۱۳
۱-٥-٤ همبستگی‌نگار ۱٤
۱-٥-٥ همسانگردی و ناهمسانگردی ۱٥
۱-٥-٦ پارامترهای تغییرنگار ۱٦
۱-٦ رابطه بین تغییرنگار و هم‌تغییرنگار ۱۷
۱-۷ پیشگویی فضایی ۱۹
۱-۸ آمار فضایی و همه‌گیری‌شناسی ۲۰
۱-۸-۱ بیماری سل ۲۱
۱-۸-۲ سیر بیماری سل ۲۲
۱-۸-۳ تشخیص بیماری سل ۲۲
۱-۸-٤ کنترل جهانی بیماری سل ۲٤
۱-۸-٥ شاخص‌های بیماری سل ریوی ۲٥
۱-۸-٦ معرفی داده‌ها ۲۷
۱-۸-۷ همه‌گیری‌شناسی بیماری سل ریوی ۲۹
۱-۹ مروری بر مطالعات انجام شده ۳۳

فصل دوم: برازش مدل تغییرنگار

۲-۱ مقدمه ۳۵
۲-۲ مدلهای تغییرنگار ۳۶
۲-۲-۱ مدلهای دارای ازاره ۳۶
۲-۲-۲ مدلهای بی ازاره ۳۹
۲-۳ مدلهای هم‌تغییرنگار ۴۰
۲-۴ برآوردهای تغییرنگار ۴۱
۲-۴-۱ برآورد گشتاوری ۴۱
۲-۴-۲ برآوردهای استوار ۴۳
۲-۵ روشهای برازش مدل تغییرنگار ۴۵
۲-۵-۱ روش ماکسیمم درستنمایی ۴۶
۲-۵-۲ روش کمترین توانهای دوم عادی ۴۷
۲-۵-۳ روش کمترین توانهای دوم تعمیم یافته ۴۷
۲-۵-۴ روش کمترین توانهای دوم موزون ۴۸
۲-۶ برازش مدل تغییرنگار به داده‌ها ۵۰
۲-۶-۱ برآورد تغییرنگار داده‌ها ۵۱
۲-۶-۲ برازش مدل تغییرنگار ۵۲
۲-۷ تعداد بهینه فاصله در برآورد تغییرنگار ۵۵
۲-۷-۱ روشهای شبیه‌سازی داده‌های فضایی ۵۶
۲-۷-۲ روش تجزیه چولسکی ۵۸
۲-۷-۳ نتیجه شبیه‌سازی ۵۹

فصل سوم: کریگیدن

۳-۱ مقدمه ۶۱
۳-۲ کریگیدن عادی ۶۳
۳-۲-۱ کریگیدن عادی بر حسب هم‌تغییرنگار ۶۵
۳-۳ کریگیدن عام ۶۸
۳-۳-۱ کریگیدن عام بر حسب هم‌تغییرنگار ۷۱
۳-۳-۲ مقایسه تغییرنگار با هم‌تغییرنگار در کریگیدن ۷۱
۳-۴ کریگیدن بلوکی ۷۲

۳-۵ هم‌کریگیدن ۷۳
۳-۶ خطای اندازه‌گیری و اثر قطعه‌ای ۷۴
۳-۷ نامانایی ۷۹
۳-۷-۱ روند زدایی داده‌ها ۸۰
۳-۷-۲ تابع کواریانس تعمیم یافته ۸۲
۳-۸ پیشگویی فضایی میزان بیماریها ۸۴
۳-۸-۱ مدل تغییرنگار داده‌ها ۸۵
۳-۸-۲ محاسبات کریگیدن ۸۶
۳-۸-۳ اعتبار سنجی متقابل تغییرنگار ۸۷

فصل چهارم: داده‌های دورافتاده فضایی

۴-۱ مقدمه ۹۰
۴-۲ روشهای شناسایی داده‌های دورافتاده فضایی ۹۲
۴-۲-۱ نمودار پراکنش h ۹۳
۴-۲-۲ ابر تغییرنگار ۹۴
۴-۲-۳ نمودارهای جعبه‌ای ابر تغییرنگار ۹۶
۴-۲-۴ استفاده از باقیمانده‌های پیشگویی ۹۸
۴-۲-۵ روش کرسی برای داده‌های مشبکه ۹۹
۴-۲-۶ روش هاینینگ ۱۰۰
۴-۲-۷ مروری برسایر روشها ۱۰۰
۴-۳ نحوهٔ برخورد با داده‌های دورافتاده ۱۰۱
۴-۴ برآوردگرهای استوار ۱۰۳
۴-۴-۱ پیشگویی استوار ۱۰۴
۴-۵ مثال کاربردی ۱۰۵

فصل پنجم: ساختن مدلهای ناهمسانگرد

۵-۱ مقدمه ۱۱۴
۵-۲ ناهمسانگردی هندسی ۱۱۶
۵-۳ تعمیم ناهمسانگردی بیضوی ۱۲۱
۵-۴ ناهمسانگردی ازاره ۱۲۳
۵-۵ ناهمسانگردی مدل ۱۲۷

۵-۵-۱ شبیه‌سازی ۱۳۶
۵-۶ مثال کاربردی ۱۴۰
۵-۶-۱ برازش مدل به تغییرنگارهای سودار داده‌ها ۱۴۰
۵-۶-۲ ساختن مدل تغییرنگار ناهمسانگرد داده‌ها ۱۴۲
۵-۶-۳ پیشگویی در موقعیتهای جدید ۱۴۳
۵-۶-۴ اعتبار سنجی متقابل مدل برازانده شده به داده‌ها ۱۴۳
۵-۷ نتیجه‌گیری ۱۴۵

فصل ششم: تهیه نقشه آماری بیماری سل ریوی

۶-۱ مقدمه ۱۴۷
۶-۲ تعیین ساختار همبستگی داده‌ها ۱۴۹
۶-۳ تهیه نقشه آماری بیماری سل ریوی ۱۵۰
۶-۴ بحث و نتیجه‌گیری ۱۵۵
۶-۵ پیشنهادها ۱۵۹

پیوست‌ها

پیوست «الف»: نام و طول و عرض مراکز شهرستانها، و میزان بروز سل ریوی ... ۱۶۳
***پیوست «ب»**: معرفی نرم‌افزارهای R و GeoR* ۱۶۵
***پیوست «ج»**: برنامه‌ها* ۱۶۹
مراجع ۱۸۸

فهرست شکل‌ها

شماره	عنوان شکل	صفحه
۱-۱	نمایش پارامترهای یک مدل تغییرنگار	۱۶
۱-۲	رابطه بین تغییرنگار و هم‌تغییرنگار بر حسب فاصله	۱۷
۱-۳	نقشه موقعیت داده‌ها	۳۱
۱-٤	نمودار توزیع فراوانی داده‌های سل ۱۳۷۷ و ۱۳۷۸	۳۲
۲-۱	نمودار مدلهای شبه تغییرنگار دارای ازاره: کروی، مکعبی، موجی، نمایی و ...	٤۱
۲-۲	برآوردهای کلاسیک و استوار شبه تغییرنگار، و مدلهای گاوسی و توانی ...	٥٤
۲-۳	ناحیه تحملِ مستطیل شکل در رأس بردار h	٥٥
۲-٤	ناحیه تحمل به صورت قطاع، برای برآورد تغییرنگار سودار	٥٦
۲-٥	نمودار RMISE در برابر تعداد فاصله (k)	٦۰
۳-۱	نمودار مقادیر پیشگویی $\hat{Z}(t_0)$ در چند نقطه بین و خارج ...	۷۹
۳-۲	نمودار تغییرنگار داده‌های اصلی و داده‌های روند زدوده، ...	۸۲
٤-۱	نمودار ابر شبه‌تغییرنگار داده‌های سل، ۱۳۷۸ بر حسب فاصله	۹٥
٤-۲	نمودار جعبه‌ای ابر شبه‌تغییرنگار داده‌های سل، ۱۳۷۸ در فواصل مختلف	۹۷
٤-۳	نمودار جعبه‌ای ابر شبه‌تغییرنگار داده‌های سل، ۱۳۷۸ در فواصل مختلف	۹۸
٤-٤	نمودار جعبه‌ای ابر ریشه تفاضل‌ها برای داده‌های سل، ۱۳۷۸ در فواصل‌مختلف	۱۰٤
٤-٥	نمودار پراکنش سه‌بعدی داده‌های سل، ۱۳۷۷	۱۰٦
٤-٦	نمودار دایره‌ای داده‌های سل، ۱۳۷۷	۱۰٦
٤-۷	نمودار پراکنش سه‌بعدی داده‌های سل، ۱۳۷۸	۱۰۷
٤-۸	نمودار دایره‌ای داده‌های سل، ۱۳۷۸	۱۰۷
٤-۹	نمودار ابر شبه تغییرنگار داده‌های سل، ۱۳۷۸، پس از حذف سه داده دورافتاده	۱۰۸
٤-۱۰	نمودار مقادیر داده‌های سل ۱۳۷۷ در برابر مقادیر پیشگویی	۱۱۰
٤-۱۱	نقشه مقادیر باقیمانده‌های پیشگویی استاندارد شده، ۱۳۷۷	۱۱۰
٤-۱۲	نمودار مقادیر داده‌های سل ۱۳۷۸ در برابر مقادیر پیشگویی	۱۱۱
٤-۱۳	نقشه مقادیر باقیمانده‌های پیشگویی استاندارد شده، ۱۳۸۷	۱۱۱
٤-۱٤	نمودار ساقه و برگ باقیمانده‌های پیشگویی استاندارد شده، ۱۳۷۷	۱۱۲
٤-۱٥	نمودار ساقه و برگ باقیمانده‌های پیشگویی استاندارد شده، ۱۳۷۷	۱۱۲
٥-۱	نمودار شبه تغییرنگارهای سودار داده‌های سل، ۱۳۷۷	۱۱٥

۵-۲ نمایش رابطه مختصات کارتزین با قطبی در فضای دو بعدی ۱۱۷
۵-۳ نمودار شبه‌تغییرنگار خطی ناهمسانگرد به ازای ۱۱۸
۵-٤ نمودار شبه‌تغییرنگار خطی ناهمسانگرد به ازای ۱۱۹
۵-۵ نمودار شبه‌تغییرنگار نمایی ناهمسانگرد به ازای ۱۲۰
۵-٦ مدلسازی بُرد به کمک برآورد بُرد در چهار سوی اصلی ۱۲۲
۵-۷ مدلسازی بُرد به کمک برآورد بُرد در چهار سوی دلخواه ۱۲۲
۵-۸ مدلسازی بُرد به کمک دو بیضی ۱۲۳
۵-۹ نمودارهای شبه‌تغییرنگار نمایی در سوی ۰ با $c_1 = ۲۴۰$ و ۱۲۵
۵-۱۰ نمودار شبه‌تغییرنگار نمایی ناهمسانگرد مدل (۵-۱۰) ۱۲۵
۵-۱۱ نمودار شبه‌تغییرنگار نمایی ناهمسانگرد مدل (۵-۱۰) با $\varphi = \pi/۴$ و اثرقطعه‌ای ۱۲۷
۵-۱۲ نمایش مجموعه نقاط دارای طول یکسان $r_{x,y}$ ۱۲۸
۵-۱۳ شبه تغییرنگار نمایی ۷۰ با ۱۳۰
۵-۱٤ نمودار شبه‌تغییرنگار نمایی ناهمسانگر(۵-۱۲) متناظر با تغییرنگارهای سودار ... ۱۳۰
۵-۱۵ شبه‌تغییرنگار کروی ۷۰ با ۱۳۱
۵-۱٦ نمودار شبه‌تغییرنگار ناهمسانگرد (۵-۱۲) متناظر با ۷۰ و ۱۳۲
۵-۱۷ نمودارهای دوران محورها به اندازهٔ φ رادیان (درجه) ۱۳٤
۵-۱۸ بردارهای $\boldsymbol{h}_1$ و $\boldsymbol{h}_2$ در سوهای ۱۳۵
۵-۱۹ برآورد کلاسیک شبه‌تغییرنگارهای سودار و مدل گاوسی برازانده شده به آن . ۱۳۷
٦-۱ نمودار شبه‌تغییرنگار گاوسی برحسب فاصله، پس از حذف داده‌های دورافتاده ۱۵۰
٦-۲ نقشه پیشگویی میزان بروز سل ریوی اسمیر مثبت، ۱۳۷۷ . . . . . ۱۵۲
٦-۳ نقشه انحراف معیار پیشگویی میزان بروز سل ریوی اسمیر مثبت، ۱۳۷۷ . ۱۵۲
٦-٤ نقشه پیشگویی میزان بروز سل ریوی اسمیر مثبت، ۱۳۷۸ . . . . . ۱۵۳
٦-۵ نقشه انحراف معیار پیشگویی میزان بروز سل ریوی اسمیر مثبت، ۱۳۷۸ . . ۱۵۳
٦-٦ نقشه کاهش میزان بروز سل ریوی اسمیر مثبت، از ۱۳۷۷ به ۱۳۷۸ . . . ۱۵٤

فصل اول

مقدمه و کلیات

۱-۱ تاریخچه

آمـار فضـایی شاخهٔ نسبتاً جدیدی از آمار است که از حدود ۳۰ سال قبل با روشهای محدودی برای تجـزیه و تحـلیل داده‌هـایی کـه با موقعیت فضایی یا مکانی همراه هستند، شروع شده و در سالهای اخیر مبانی نظری و کاربردی آن مورد توجه متخصصین آمار قرار گرفته است. همانطور که در بخش بعـد مـی‌آیـد، مهمـترین ویژگـی ایـن داده‌هـا، کـه تجـزیه و تحلیل آنها را تحت تاثیر قرار میدهد، همبسـتگی آنهاسـت. وجـود همبسـتگی فضـایی در داده‌هـا یـک مسأله قدیمی‌تر بوده که بعضی از آمارشناسـان مـتوجه آن شـده‌انـد منـتها قادر به مدلسازی و استفاده از آن در تجزیه و تحلیل داده‌ها نـبوده‌انـد. فیشـر از جملـه افـرادی بـود کـه مـتوجه وجـود همبستگی فضایی در طرح آزمایشهای کشـاورزی شده بود. وی در مقالهٔ سال ۱۹۳۵ خود نوشت: «بعد از انتخاب یک ناحیه، ما هیچ دلیلی بـرای ایـن حقیقـت مهـم نداریم که چرا کرت‌های نزدیک به هم از لحاظ محصول شباهت بیشتری نسـبت بـه کـرت‌هـای دور از هم دارند». وی بر سر یک دوراهی قرار داشت: آیا این تغییرپذیری را مدلسـازی کـند یـا آن را حـذف کـند. وی تصـمیم گرفت با تصادفی‌سازی و زیاد کردن فاصله بین کـرت‌هـا، اریبی ناشی از این همبستگی را در برآورد اثر تیمارها حذف کرده یا کاهش دهد. چنانچه ییـتس (۱۹۳۸) نیز بیان می‌کند «تصادفی‌سازی علاوه بر کنترل اریبی ناخواسته، اثر همبستگی فضایی را نیز کاهش می‌دهد، لیکن آن را کاملاً از بین نمی‌برد».

اسـمیت (۱۹۳۸) نشـان داد هر چه اندازهٔ کرت بزرگتر شود، واریانس خطا به میزان کمی کاهش مـی‌یابد. با وجود تجربی بودن تحلیل وی، بیان دقیق مسأله، وجود همبستگی فضایی در آزمایشهای کشاورزی را القا می‌کند. اما مدلهای آماری برای تبیین چنین پدیده‌هایی تا سالها بعد نتوانستند ظهور پیدا کنند (ویتل، ۱۹۵٤). روشهای نزدیکترین همسایه با استفاده از اختلاف کرتهای مجاور به عنوان مـتغیر کمکـی، همبسـتگی فضـایی را بـه طـور غیـر مستقیم در تجزیه و تحلیل آزمایشهای میدانی کشاورزی وارد می‌کرد. در بخش ۵-۷-۱ کِرسی (۱۹۹۳) می‌توان مختصری از این روشها را ملاحظه نمود.

امـروزه تجـزیه و تحـلیل داده‌هـای فضـایی در شـاخه‌هـای مختـلف علوم از قبیل زمین‌شناسی، کشـاورزی، پزشـکی، بـوم‌شناسی، محیط زیست و دورسنجی به یکی از مباحث مهم آمار بدل شده اسـت. اگـرچه تجـزیه و تحـلیل و گسترش دانش آمار فضایی عمدتاً به وسیلهٔ آمارشناسان صورت مـی‌گیـرد، امـا پیدایش و کاربرد عملی این شاخه از آمـار به وسیـلهٔ زمین‌شنـاسـان و تحت عنـوان زمیـن آمـار[1] صـورت گـرفت و کـتابها و مقـالات مختـلفی نیز در این زمینه چاپ شده و هنوز هم مـی‌شـود. از ایـن رو عجیـب نیسـت کـه در زمیـن آمـار بعضاً اصطلاحاتی به کار می‌رود که برای آمارشناسان کلاسیک ناآشنا است. هوپر[2] و واترمایر[3] نخستین افرادی بودند که از روشهای آماری در زمیـن‌شناسـی اسـتفاده کـردند و با مطالعات خود الگوهای توزیع فراوانی طلا را در معادن آفریقای جـنوبی شناسـایی کـردند. بعدها تروسکات[4] (۱۹۲۹) روشی را برای محاسبه برآورد میانگین ذخیرهٔ طـلا ارائه کرد. سیشل[1] (۱۹٤۷) مدلی را برای توزیع طلا معرفی کرد و مسأله خطاهای سیستماتیک در نمونـه‌گیـری را مطرح ساخت. وی همچنین فرمول و جدولی را به منظور محاسبه دقت میانگین موضعی متغیرها و بازهٔ اطمینان آن ارائه کرد. روش وی به گونه‌ای بود که از موقعیت فضـایی داده‌ها اسـتفـاده‌ای نمـی شـد، برای رفع این کاستی درسالهای ۱۹۵۰ و ۱۹۵۱ مطالعات تکمیلی گسترده‌ای

---

[1] Geostatistics

[2] Hooper

[3] Watermeyer

[4] Troscott

توسط افرادی چون راس[۲]، کریگ[۳]، و دویس[۴] صورت گرفت (مدنی، ۱۳۷۳). مدتی بعد با مشاهدهٔ نوعی رابطه بین عیار یک کانی در بخشهای مختلف یک منطقه معدنی، تلاشهایی برای یافتن ارتباط فضایی نمونه‌ها شروع شد که از جمله می‌توان به کارهای ویتن[۵]، کرومباین[۶]، گریفیت[۷]، کوچ[۸]، لینک[۹] و هاربو[۱۰] اشاره کرد. این کوششها در دههٔ ۱۹۵۰ و اوایل دههٔ ۱۹۶۰ به ابداع و تکمیل روش «تجزیه و تحلیل سطوح روند» منجر شد. ماترون در دههٔ ۱۹۶۰ گام بعدی را برای گسترش استفاده از روشهای آماری در علوم زمین برداشت و با انتشار مقاله ای در سال ۱۹۶۲ پایه‌های شاخه زمین آمار را بنا نهاد. وی در سال ۱۹۶۵ مبحث متغیرهای ناحیه‌ای و در سال ۱۹۶۹ بحث کریگیدن را، که منتسب به دی. جی. کریگ، مهندس معدن افریقای جنوبی است، مطرح ساخت. دیوید (۱۹۷۷) در کتاب «برآورد زمین آماری ذخیرهٔ طلا» و ژورنل و هوبرگتز (۱۹۷۸) در کتاب «زمین آمار در معدن» روشهای زمین آمار را با اصطلاحات خاص شرح دادند. ریپلی (۱۹۸۱) دریچه دیگری را به زمین آمار گشود و در آن از اصطلاحات آماری استفاده کرد. برای این منظور وی نظرات مربوط به پیشگویی میدانهای تصادفی را که ارتباط نزدیکی با سریهای زمانی دارند، به کار گرفت.

از واژهٔ «زمین آمار» برداشتهای مختلفی شده است. زمین آمار در مفهوم اروپایی آن به شاخه‌ای از علم آمار گفته می‌شود که مبتنی بر «نظریه متغیرهای ناحیه‌ای» (ماترون، ۱۹۶۵) است و با متغیرها یا داده‌های فضایی سروکار دارد، از این رو با «آمار فضایی» مترادف است. اما در مفهوم آمریکایی، به کاربرد تمام روشهای آماری در علوم زمین از جمله آمار کلاسیک و آمار فضایی اطلاق می‌شود. این

---

[1] Sichel

[2] Ross

[3] Krige

[4] Dewijs

[5] Whitten

[6] Krumbein

[7] Greffit

[8] Koch

[9] Link

[10] Harbugh