

روشهای پیشرفته آماری در تحلیل داده‌های باستان‌شناختی

کمال‌الدین نیکنامی

تهران

۱۳۹۰



سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)

مرکز تحقیق و توسعه علوم انسانی

سرشناسه: نیکنامی، کمال‌الدین، ۱۳۳۵.
عنوان و نام پدیدآور: روشهای پیشرفته آماری در تحلیل داده‌های باستان‌شناختی / کمال‌الدین نیکنامی.
مشخصات نشر: تهران: سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)، مرکز تحقیق و توسعه علوم انسانی، ۱۳۹۰.
مشخصات ظاهری: دوازده، ۵۵۹ ص.؛ مصور، جدول، نمودار.
نروست: «سمت»؛ ۱۵۳۵. باستان‌شناسی؛ ۵۱.
شابک: ۹۷۸-۹۶۴-۵۳۰-۶۹۴-۴ ۸۳۰۰۰ ریال
وضعیت فهرست‌نویسی: فیا
یادداشت: پشت جلد به انگلیسی: Kamal Aldin Niknami. Advanced Quantitative Methods in Archaeological Data Analysis.
یادداشت: ویژه نامه.
یادداشت: کتابنامه: ص ۵۴۶-۵۵۹.
موضوع: باستان‌شناسی — روشهای آماری
شناسه افزوده: سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)، مرکز تحقیق و توسعه علوم انسانی.
رده‌بندی کنگره: ۱۳۹۰ ۹۷۸/۵۳۰/۶۹۴/۴
رده‌بندی دیویی: ۹۳۰/۱۰۲۸
شماره کتابشناسی ملی: ۲۶۳۳۷۲۴

سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)



مرکز تحقیق و توسعه علوم انسانی

روشهای پیشرفته آماری در تحلیل داده‌های باستان‌شناختی

دکتر کمال‌الدین نیکنامی (استاد دانشگاه تهران)

چاپ اول: زمستان ۱۳۹۰

تعداد: ۱۰۰۰

حروفچینی و لیتوگرافی: «سمت»

چاپ و صحافی: سازمان چاپ و انتشارات اوقاف

قیمت: ۸۳۰۰۰ ریال. در این قیمت چاپ مذکور ثابت است و فروشنده‌گان و عوامل توزیع مجاز به تغییر آن نیستند.

نشانی ساختمان مرکزی: تهران، بزرگراه جلال آل احمد، غرب پل یادگار امام (ره)، روبه‌روی بعب گاز، کد پستی ۱۶۶۳۶، تلفن ۰۲۱-۴۴۲۴۶۲۵.

www.samt.ac.ir

info@samt.ac.ir

هر شخص حقیقی یا حقوقی که تمام یا قسمتی از این اثر را بدون اجازه ناشر یا مؤلف، نشر یا پخش یا عرضه یا تکثیر یا تجدید چاپ نماید، مورد پیگرد قانونی قرار خواهد گرفت.

سخن «سمت»

یکی از اهداف مهم انقلاب فرهنگی، ایجاد دگرگونی اساسی در دروس علوم انسانی دانشگاهها بوده است و این امر، مستلزم بازنگری منابع درسی موجود و تدوین منابع مبنایی و علمی معتبر و مستند با در نظر گرفتن دیدگاه اسلامی در مبنای و مسائل این علوم است.

ستاد انقلاب فرهنگی در این زمینه گامهایی برداشته بود. اما اهمیت موضوع اقتضا می کرد که سازمانی مخصوص این کار تأسیس شود و شورای عالی انقلاب فرهنگی در تاریخ ۶۳/۱۲/۷۱ تأسیس «سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها» را که به اختصار «سمت» نامیده می شود، تصویب کرد.

بنابراین، هدف سازمان این است که با استمداد از عنایت خداوند و همت و همکاری دانشمندان و استادان متعهد و دلسوز، به مطالعات و تحقیقات لازم بپردازد و در هر کدام از رشته های علوم انسانی به تألیف و ترجمه منابع درسی اصلی، فرعی و جنبی اقدام کند.

دشواری چنین کاری بر دانشمندان و صاحب نظران پوشیده نیست و به همین جهت مرحله کمال مطلوب آن، باید به تدریج و پس از انتقادهای و یادآوریهای پیاپی ارباب نظر به دست آید و انتظار دارد که این بزرگواران از این همکاری دریغ نورزند. کتاب حاضر برای دانشجویان رشته باستان شناسی، در مقطع کارشناسی به عنوان منبع اصلی درس «روشهای تحلیل آماری» به ارزش ۲ واحد و در مقطع کارشناسی ارشد به عنوان منبع اصلی دروس «روش شناسی تحقیق» و «آمار در باستان شناسی» به ارزش ۴ واحد تدوین شده است. امید است که علاوه بر جامعه دانشگاهی، سایر علاقه مندان نیز از آن بهره مند شوند.

از استادان و صاحب نظران ارجمند تقاضا می شود با همکاری، راهنمایی و پیشنهادهای اصلاحی خود، این سازمان را در جهت اصلاح کتاب حاضر و تدوین دیگر آثار مورد نیاز جامعه دانشگاهی جمهوری اسلامی ایران یاری دهند.

فهرست مطالب

صفحه	عنوان
هشت	مقدمه مؤلف
۱	فصل اول: مقدمه تحلیل چندمتغیری آماری
۸	تعریف و طبقه‌بندی متغیرها
۱۴	فرایند آماده‌سازی داده‌ها برای آنالیز
۱۸	تکنیکهای وارد کردن داده‌ها
۲۶	تشکیلات داده‌ها
۳۰	غریبال کردن داده‌ها و تبدیل داده‌ها
۳۵	انتخاب مناسب‌ترین روش تبدیل داده‌ها
۳۷	ارزیابی تقارن با استفاده از ترسیم جعبه و سکه (پلات ویسکر)
۳۹	ارزیابی نرمال بودن با ترسیم منحنی احتمالات نرمال یا چندک نرمال
۴۳	انتخاب روش تبدیل داده برای ایجاد حالت نرمال
۴۶	سنجشهای آماری برای ارزیابی نرمال بودن
۴۸	ارزیابی نیاز به تبدیل داده
۴۹	ارزیابی استقلال داده‌ها
۵۲	انتخاب آنالیزهای مناسب
۵۳	اندازه‌های آماری مناسب
۵۴	مقیاسهای متغیر اسمی
۵۵	مقیاس داده‌های ترتیبی
۵۶	مقیاس داده‌های فاصله‌ای
۵۶	مقیاس داده‌های نسبتی
۵۷	انتخاب مناسب‌ترین آنالیزهای چندمتغیری

۶۱	فصل دوم: ماتریسهای جبری
۶۶	عملیات چهار عمل اصلی بر روی ماتریسها
۶۶	جمع و تفریق ماتریسها
۶۸	ضرب ماتریسها
۷۲	شرایط ضرب پذیری دو ماتریس
۷۳	دارون ماتریس
۸۰	فصل سوم: رگرسیون چندگانه
۸۳	تخمین و سنجش رگرسیون چندگانه
۸۹	ارتفاع و فاصله تا منابع آب در متغیر فراوانی سفال
۱۳۴	فصل چهارم: رگرسیون لجستیک
۱۴۴	محاسبه مدل رگرسیون لجستیک
۱۴۷	محاسبه نیکویی برازش
۱۵۰	محاسبه ASE برای لجیت (p)
۱۵۳	رگرسیون لجستیک چندمتغیری
۱۷۱	فصل پنجم: مدل لوگ (لگاریتم) خطی
۱۷۳	تقسیم بندی فرایند لوگ خطی
۱۸۸	فصل ششم: تحلیل مؤلفه های اصلی
۱۸۸	مفاهیم اساسی آنالیز مؤلفه های اصلی
۱۹۴	فرایندهای اجرای آنالیز مؤلفه های اصلی
۲۴۳	فصل هفتم: تحلیل تشخیصی
۲۵۵	انواع طبقه بندی و قواعد آنها
۲۵۹	تابع تشخیص کانونی و تابع گروه بندی فیشر
۳۰۴	تعیین صحت گروه بندی
۳۱۴	فصل هشتم: تحلیل خوشه ای
۳۲۶	انواع خوشه بندی

صفحه	عنوان
۳۳۱	تعیین فاصله میان خوشه‌ها
۳۴۸	روشهای خوشه‌بندی مقسم (تقسیمی)
۳۵۴	روش خوشه‌بندی تفکیکی
۳۸۷	فصل نهم: تحلیل تناظری
۳۹۸	فاصله خن دو (k^2) و اینرشیا در آنالیز CA
۴۰۵	ترسیم فاصله χ^2
۴۶۳	واژه‌نامه توصیفی موضوعی
۵۴۶	منابع

مقدمه مؤلف

اولین هدف این کتاب معرفی روشهای آماری پیشرفته برای باستان‌شناسان است. گرچه مثالهایی که برای کاربرد قواعد و روشهای پیشرفته، همگی دارای موضوعات باستان‌شناسی اند ولی این بدین معنا نیست که سایر رشته‌های مرتبط نتوانند از آن استفاده کنند. امروزه باستان‌شناسی با بسیاری از علوم، روابط تحقیقاتی میان‌رشته‌ای ایجاد نموده و علاوه بر استفاده از بعضی مفاهیم نظری آن رشته‌ها، به صورت گسترده‌ای از روشهای عملیاتی آنها نیز سود می‌برد که از آن جمله می‌توان به روشهای جغرافیایی، انسان‌شناسی و به خصوص انسان‌شناسی فیزیکی و زیستی، زیست‌شناسی و فیزیولوژی، زمین‌شناسی، جامعه‌شناسی، علوم کامپیوتر و تعداد زیاد دیگری از علوم تجربی، علوم اجتماعی و انسانی و حتی هنر را می‌توان نام برد. امروزه، تحقیقات علمی باستان‌شناسی در همه جهان یک سیمای دوگانه یافته است. یک سیمای آن که با تأثیر از مطالعات میان‌رشته‌ای قابل شناسایی است بیشتر دارای صفات روش‌شناسیهای علوم تجربی است و مبنای آن را عمدتاً روشهای علوم تجربی و بیشتر از همه راهبردهای تحلیلیهای کمی تشکیل می‌دهند. این گرایشها بیشتر از همه، از این اندیشه‌ها سرچشمه می‌گیرند که اولاً اغلب داده‌های باستان‌شناختی در ذات خود دارای صفات کمی‌اند و ثانیاً، بسیاری از سؤالاتی هم که امروزه در حوزه علوم باستان‌شناختی مطرح‌اند فقط با روشهای علمی قابل پاسخگویی هستند. این اندیشه‌ها و اندیشه‌های مشابه، یک تغییر رویکرد محسوسی را از نیمه‌های قرن گذشته میلادی در تحقیقات باستان‌شناسی باعث شده‌اند که به لحاظ روش‌شناسی، با سیمای دیگر.

تحقیقات باستان‌شناختی که اغلب مبتنی بر روشهای توصیفی و استنتاجی قیاسی است، متفاوت است. اکنون در جهان هر کدام از این رویکردهای تحقیقاتی، طرفداران خاص خود را دارند و هر کدام بر ارزشهای رویکردهای خود اصرار می‌ورزند. در اینجا قصد بر این نیست که نقاط ضعف و قوت این رویکردها مورد نقد و بررسی قرار گیرند بلکه بیشتر تأکید بر این است که یک راه و رسم دیگر اندیشیدن و استدلال را برای جامعه باستان‌شناسی معرفی نماید. در باستان‌شناسی استدلال بر اساس تحلیل داده‌های کمی که منجر به اقتباس و یا ابداع روشهای مرتبط شد، نسبت به سایر گرایشهای علمی سابقه بس طولانی ندارد و به طور عمده پس از آغاز انقلاب فکری باستان‌شناسی در نیمه‌های قرن بیستم میلادی، توانسته است اکثر حوزه‌های تحقیقاتی باستان‌شناسی را تحت تأثیر قرار دهد. خوشبختانه در ایران هر چند بسیار دیر ولی اکنون باستان‌شناسان رغبت فراوانی را برای این نوع شیوه استدلال از خود نشان می‌دهند و به خصوص در جایی که کسی درگیر تحقیقات میان‌رشته‌ای در باستان‌شناسی است این ضرورت را به صورت اجتناب‌ناپذیری احساس می‌کند. اما برای باستان‌شناسان مشکل دیگری وجود داشت و آن اینکه اکثریت آنان دارای سوابق تحصیلات علوم انسانی و متبخر در روشهای استدلال نظری اند و کسب مهارت در روشها و مدل‌های استدلالی ریاضی ممکن است برای آنان یک تمایل دست‌نیافتنی محسوب گردد. از طرف دیگر نبود واحدهای درسی مرتبط با موضوع در سرفصل‌های دانشگاهی این رشته و همچنین نبود منابع آموزشی مرتبط نیز به این مشکلات دامن می‌زنند. بنابراین اولین قدم در پاسخگویی به این نیاز می‌توانست تدوین یک سری منابعی باشد که ضمن بیان ارزشهای علمی روشهای تحلیل و استدلال کمی، بتواند مهارتها و قابلیت‌های اجرایی را برای کاربران منتقل نماید. به همین خاطر، هم در بخش مقدماتی این کتاب، با نام روشهای تحلیل کمی در پژوهشهای باستان‌شناسی (۱۳۸۷)، و هم در این مجموعه که به روشهای پیشرفته می‌پردازد، تا حد امکان سعی شده است:

۱. از میان تعداد کثیری از روشها، بیشتر به آنهایی پرداخته شده است که در

فرایند اجرایی خود دارای پیچیدگیهای کمتری اند و اغلب روشهایی اند که در ادبیات باستان‌شناسی آشنایی بیشتری با آنها وجود دارد.

۲. سعی شده است در بیان هر کدام از روشها یک ساختار دوبخشی کاملاً رعایت گردد بدین معنی که ابتدا اصول و قواعد ریاضی هر کدام از آنها بدون پرداختن به جزئیات بیشتر بیان شده و سپس کاربرد آن در باستان‌شناسی به طور مفصلی بحث شده است. در بسیاری موارد از مثالهای کاملی استفاده شده و فرایند تحلیل از ابتدا تا تفسیر نشان داده شده است. دانشجویانی که با ملاحظه بخش نخستین هر کدام از فصول و با مشاهده فرمولها، ممکن است یک تلقی نگران‌کننده‌ای نسبت به ادامه بحث پیدا کنند باید این بخشها را به حال خود بگذارند. در طراحی تشریح فرایند هر روش سعی شده است سابقه دانش ریاضی خوانندگان مورد نظر قرار نگیرد و بحثها عموماً طوری تداوم دارند که کسانی هم که با دانش ریاضی کمترین ارتباطی دارند بتوانند این روشها را به طور مستقل به کار گیرند.

۳. گرچه بسیار سعی شده است این کتاب یک کتاب آشپزی نباشد ولی گاهی به اجبار به آن طرف سویی داده شده است. عمده‌ترین دلایل آن را در ابتدای این مقدمه ذکر کردم. ولی آن چیزی که بعد از انتشار بخش مقدماتی با آن روبه‌رو شدم درخواستهای مکرر خوانندگان و دانشجویانی بود که شرح اجرای کامپیوتری این روشها را نیز در ادامه هر فصل خواستار بودند. این کار را به این دلیل انجام ندادم زیرا انجام روشها بدون درک مفاهیم آن، گمراهیهای بزرگی را در استنباط سبب می‌شود. به این مطلب با توضیح بیشتری در فصل اول همین مجموعه پرداخته شده است.

۴. در بخش منابع، علاوه بر منابع مورد استفاده منابع مرتبط دیگری نیز آورده شده‌اند که یک قلمرو وسیعی از کاربردهای روشهای کمی را بر روی انواع داده‌های باستان‌شناختی برای خوانندگان فراهم کرده است. سعی شده است منابع مطالعاتی به تفکیک روشها آورده شوند و اغلب آنها را آنالیزهایی تشکیل می‌دهند که در سالهای بسیار اخیر در تشریحات علمی ظاهر شده‌اند. این کتاب با ذکر منابع مطالعاتی، علاوه بر آشنا کردن خوانندگان با ادبیات وسیع رشته خود، مزیت

دیگری را هم تعقیب می کند و آن اینکه به لحاظ گستردگی و فشردگی مطالب امکان شرح و تفسیر عملیاتهای انجام گرفته میسر نبوده است و خوانندگان می توانند با رجوع به آن منابع علاوه بر کاربرد روشهای مورد نظر خود، با روند استدلال و تفاسیر نهایی نیز آشنا گردند.

در تدوین این کتاب افراد زیادی سهم اند و بی تردید، بدون همکاری آنها این مجموعه هیچ وقت شانس انتشار پیدا نمی کرد. نخستین سهم مربوط به سازمان سمت است که فقط او می تواند این گونه نوشته های مبانی را که خریداران چندانی هم ندارند، پشتیبانی کند و این همان نکته ای است که سازمان سمت را در راستای یکی از اهدافش (سرساری برای تولید و رشد علم) بسیار موفق نموده است. سازمان سمت و به خصوص گروه پژوهشی باستان شناسی آن سازمان، بیش از حد وظایف رسمی خود برای انتشار آن تلاش نمودند که بدین وسیله تشکر قلبی خود را اعلام می دارم. همچنین، از سرکار خانم زهرا عطارد، که ویراستاری علمی این کتاب به عهده ایشان بود و با حوصله فراوان توانستند اشتباهات فاحش کتاب را تصحیح کنند، تشکر می کنم.

دو نفر دیگر نیز سهم فراوانی در تدوین این کتاب دارند. یکی سرکار خانم دکتر مژگان جایز که با مهربانی هرچه تمام تر کلیه نمودار و نقشه های کتاب را بازآفرینی و تنظیم نمود و دیگری سرکار خانم دکتر مریم رضایی که تنظیم تمامی منابع را به انجام رسانید، در این نوشته اینجانب خود را مدیون آنها می دانم.

در پایان اشاره به این نکته هم لازم است گرچه این کتاب عنوان روشهای پیشرفته دارد ولی هنوز جزئی از روشهای کلاسیک محسوب می گردد و تنها ارزشی که می توان به آن قائل شد این است که توانسته است در یک فرایند روشمند، اصلی ترین روشهای مورد نیاز را به زبان ساده ای در یک جا جمع نماید ولی نباید فراموش کرد که جهان پردازی امروزه به شدت در حال تحول است و اکنون متخصصان تحلیل گر داده های باستان شناختی، از روشهای بسیار فراتر از روشهای آمار کلاسیک استفاده می نمایند. یکی از گرایشهای باستان شناسی که در سالهای

بسیار اخیر تحول شگرفی یافته است پهن‌دشت باستان‌شناسی است که به نوعی جایگزین روشها و دیدگاههای سنتی بررسیهای میدانی باستان‌شناختی شده است. در این گرایش استفاده از دو رویکرد تحلیلی که اصطلاحاً تحلیل‌های فضایی و زمین آمار خوانده می‌شوند افقهای تازه‌ای را در تحلیل داده‌های سطحی، در پیش روی باستان‌شناسان قرار داده است که برخلاف روشهای کلاسیک جزئی از روشهای مدرن محسوب می‌گردند. ابداع روشهای پردازشی نوین و ورود ماشینهای پردازشگر جدید به بازار به شدت باعث ایجاد بسترهای مناسبی برای تغییر رویکردهای پیشین تحلیل گران می‌گردد. یکی از آنها که به طور مستقیم با داده‌پردازیهایی بسیار پیشرفته باستان‌شناسی در ارتباط است سیستم اطلاعات جغرافیایی یا همان GIS می‌باشد که ابزار شگرفی در پردازش داده‌های پهن‌دشت باستان‌شناختی محسوب می‌گردد. هم‌اکنون، استفاده از GIS در باستان‌شناسی ایران کم‌کم جای خود را باز می‌کند ولی هنوز در مرحله تولید نقشه‌های توصیفی است و برای رسیدن به تولید نقشه‌های تحلیلی، به شرح و انتشار و آموزش روشهای مخصوص تحلیلی این نرم‌افزار، که عمدتاً بر پایه روشهای تحلیل فضایی و زمین آمار هستند، نیاز دارد. این حاصل نخواهد شد مگر اینکه در آینده متخصصان دیگری، سلسله بحثهای روشهای پردازشی کلاسیک را در یک تداوم توسعه‌ای به بحثهای مذکور پیشرفته پیوند دهند. برای همه کسانی که در این زمینه تلاش می‌کنند از خداوند آرزوی توفیق می‌نمایم.