

محاسبات آماری با R

تألیف

محمد مرادی و حمزه جلیلیان

گروه آمار دانشگاه رازی

انتشارات علمی پارسیان

۹-۵۷-۹۴۴۸-۶۰۰-۹۷۸

۴۹۸۳۷۲۱

محاسبات آماری با R / تألیف محمد مرادی و عبدالله جلیلیان.

تهران: انتشارات علمی پاریسیان، ۱۳۹۶.

۲۹۶ ص.: مصور، جدول، نمودار.

عنوان به انگلیسی: Statistical Computing with R.

واژه‌نامه، کتاب‌نامه، نمایه.

آر (زبان برنامه‌نویسی کامپیوتر).

R (Computer programming language)

آمار ریاضی-داده‌پردازی.

Mathematical statistics-Data processing.

۵۱۹/۵۰۲۸۵۱۳۳

۱۳۹۶ ۴/۴۴/۴۵ QAY۲۷۶

مرادی، محمد، ۱۳۵۵ دی-

جلیلیان، عبدالله، ۱۳۶۲-

فپیا.

شابک

شماره‌ی کتاب‌شناسی ملی

عنوان و نام پدیدآور

مشخصات نشر

مشخصات ظاهری

یادداشت

یادداشت

موضوع

موضوع

موضوع

موضوع

رده‌بندی دیوی

رده‌بندی کنگ

سرستان

شناسه‌ی افزوده

وضعیت فهرست‌نویسی

انتشارات علمی پاریسیان

آدرس: تهران، صندوق پستی ۳۷۹۵ - ۱۶۷۶۵

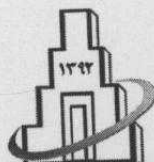
مدیر داخلی: رتولدا: مهدیه شهبازی

تلفن ثابت: ۸۸۰۶۰۳۱

تلفن همراه: ۰۹۱۷۰۱۲۲۹ و ۰۹۳۶۶۰۳۳۱۷۰

پایگاه اینترنتی: <http://parsiensub.ir>

پیام‌نگار: ma@parsiensub.ir



انتشارات علمی پاریسیان

۲۰۱

نام کتاب

مؤلفان

ویراستار

حروف‌نگار و صفحه‌آرا

طراح جلد

ناشر

نوبت چاپ

شمارگان

شابک

قیمت

محاسبات آماری با R

محمد مرادی و عبدالله جلیلیان

مهدیه شهبازی

عبدالله جلیلیان

ناصر مجلسی‌راد

انتشارات علمی پاریسیان

اول، زمستان ۱۳۹۶

جلد ۶۰۰

۹-۷-۹۴۴۸۵-۶۰۰-۹۷۸

۳۰,۰۰۰ تومان

حق چاپ برای انتشارات علمی پاریسیان محفوظ است.

فهرست مطالب

آ	پیش‌گفتار
۱	۱ مقدمه
۲	۱.۱ تاریخچه‌ی R
۳	۲.۱ دریافت و نصب نرم افزار
۵	۳.۱ محیط R
۸	۱.۳.۱ نوشتن دستورها در یک فایل اسکریپت
۱۰	۲.۳.۱ بسته‌های نرم‌افزاری R
۱۲	۴.۱ انجام محاسبات ریاضی در R
۱۲	۱.۴.۱ به‌کارگیری R به عنوان یک ماشین حساب
۱۴	۲.۴.۱ اعداد مختلط و مقدارهای عددی ویژه در R
۱۶	۳.۴.۱ دقت محاسبات عددی در کامپیوتر
۱۸	۴.۴.۱ عبارت‌های مقایسه‌ای و منطقی در R
۱۹	۵.۱ توزیع‌های احتمال در R
۲۷	۲ ساختار داده‌ها در R
۲۷	۱.۲ بردار
۲۷	۱.۱.۲ وارد کردن مؤلفه‌های یک بردار در R
۳۴	۲.۱.۲ ویرایش و مرتب‌سازی بردار
۳۷	۳.۱.۲ محاسبات برداری

۳۹	۴.۱.۲ محاسبه‌ی آمارها
۴۲	۲.۲ ماتریس و آرایه
۴۲	۱.۲.۲ ساختن یک ماتریس در R
۴۶	۲.۲.۲ ویرایش و مرتب‌سازی ماتریس
۴۹	۳.۲.۲ محاسبات ماتریسی
۵۳	۴.۲.۲ آرایه
۵۶	۳.۲ چارچوب داده‌ای
۵۸	۴. فهرست
۶۳	۳ رسم نمودارها
۶۳	۱.۳ تابع عمومی plot
۶۴	۱.۱.۳ تعیین برجستگی محورها
۶۴	۲.۱.۳ تعیین حدود محورها
۶۵	۳.۱.۳ تعیین عنوان برای نمودار
۶۵	۴.۱.۳ تعیین علامت نمایش‌دهنده‌ی نقطه در نمودار
۶۵	۵.۱.۳ تعیین اندازه‌ی نقطه‌های نمودار
۶۶	۶.۱.۳ تعیین رنگ‌ها در نمودار
۶۶	۷.۱.۳ رسم گونه‌های مختلف نمودارها
۶۸	۸.۱.۳ رسم ماتریس‌ها و چارچوب‌های داده‌ای
۶۹	۹.۱.۳ رسم نمودار یک تابع
۷۰	۱۰.۱.۳ افزودن موارد بیش‌تر به یک نمودار رسم شده
۷۲	۲.۳ بافت‌نگار
۷۵	۳.۳ نمودار چندک-چندک
۷۶	۴.۳ نمودار جعبه‌ای
۸۰	۵.۳ نمودار دایره‌ای
۸۲	۶.۳ نمودار میله‌ای

۱۵	۴ برنامه‌نویسی
۱۵	۱.۴ تابع‌نویسی
۹۱	۲.۴ ساختارهای شرطی
۹۵	۳.۴ ساختار چندانتخابی
۹۷	۴.۴ ساختارهای تکرار
۱۰۰	۵.۴ چند مثال از تابع‌نویسی در R
۱۰۰	۶.۴ خط یک عدد در مبنای دو
۱۰۲	۷.۴ تولید عددهای تصادفی با دنباله‌ی فان در کورپوت
۱۰۵	۸.۴ آزمون z
۱۰۷	۹.۴ نمونه‌گیری بدون جای
۱۰۷	۱۰.۴ تقریب عددی به روش مونت کارلو
۱۰۸	۱۱.۴ تقریب توزیع یک اماره به روش مونت کارلو
۱۱۱	۱۲.۴ تولید عددهای تصادفی به روش تبدیل ران
۱۱۵	۱۳.۴ تولید عددهای تصادفی به روش نمونه‌گیری رد
۱۱۶	۱۴.۴ شبیه‌سازی یک زنجیر مارکوف
۱۱۸	۱۵.۴ مسأله‌ی سوزن بوفون
۱۲۰	۱۶.۴ ساختن شکل‌های متحرک
۱۲۱	۱۷.۴ اجرای دستورها در حالت پردازش دسته‌ای
۱۲۵	۵ استنباط آماری برای یک و چند جامعه
۱۲۵	۱.۵ بررسی یک جامعه
۱۲۶	۱.۱.۵ برآورد ماکسیمم درست‌نمایی پارامترهای یک جامعه
۱۲۹	۲.۱.۵ آزمون توزیع جامعه
۱۳۲	۳.۱.۵ آزمون فرض و بازه‌ی اطمینان برای میانگین یک جامعه
۱۳۵	۴.۱.۵ آزمون فرض و بازه‌ی اطمینان ناپارامتری برای مرکز یک جامعه
۱۳۷	۲.۵ مقایسه‌ی دو جامعه‌ی مستقل
۱۳۷	۱.۲.۵ آزمون فرض و بازه‌ی اطمینان برای اختلاف میانگین‌های دو جامعه

۲.۲.۵	آزمون فرض و بازه‌ی اطمینان برای نسبت واریانس‌های دو جامعه	۱۴۰
۳.۲.۵	آزمون فرض و بازه‌ی اطمینان ناپارامتری برای اختلاف پارامترهای مکان و	
	نسبت پارامترهای مقیاس دو جامعه	۱۴۱
۴.۲.۵	آزمون هم‌توزیعی دو جامعه	۱۴۲
۳.۵	مقایسه‌های جفت‌شده	۱۴۳
۴.۵	مقایسه‌ی چند جامعه‌ی مستقل	۱۴۵
۱.۴.۵	آزمون برابری میانگین‌های چند جامعه	۱۴۵
۲.۴.۵	آزمون ناپارامتری برابری مرکزهای چند جامعه	۱۴۸
۳.۴.۵	آزمون همگنی واریانس چند جامعه	۱۴۹
۵.۵	نسبت‌ها	۱۵۰
۱.۵.۵	آزمون فرض و بازه‌ی اطمینان برای نسبت در یک جامعه	۱۵۰
۲.۵.۵	آزمون برای نسبت‌های چند جامعه	۱۵۲
۶.۵	استنباط آماری بر اساس روش‌های نمونه‌گیری	۱۵۴
۱.۶.۵	بازه‌ی اطمینان چندگانه‌ی حواگردان	۱۵۴
۲.۶.۵	آزمون جایگزینی برای مقایسه‌ی میانگین دو جامعه‌ی مستقل	۱۵۶

۶ تحلیل رگرسیونی

۱۵۹	
۱.۶	مدل رگرسیون خطی
۱۶۰	
۲.۶	برآزش مدل با تابع $1m$
۱۶۲	
۱.۲.۶	داده‌های شاخص باروری سوئیس
۱۶۴	
۲.۲.۶	داده‌های نگرش سازمانی کارکنان یک سازمان مالی
۱۷۱	
۳.۶	بررسی نیکویی برآزش و ارزیابی مدل
۱۷۵	
۴.۶	پیش‌بینی یک مقدار آتی
۱۸۱	
۵.۶	انتخاب متغیر به روش گام‌به‌گام
۱۸۳	

۷ تحلیل آماری در حضور متغیرهای رسته‌ای

۱۹۱	
۱.۷	جدول‌های توافقی
۱۹۱	

۱۹۱	جدول‌های توافقی دوطرفه	۱.۱.۷
۱۹۵	آزمون استقلال	۲.۱.۷
۱۹۶	آزمون همگنی	۳.۱.۷
۱۹۷	جدول‌های توافقی چندبعدی	۴.۱.۷
۲۰۰	مدل لگ خطی برای جدول‌های توافقی	۵.۱.۷
۲۰۲	طرح آزمایش‌های یک‌عاملی	۲.۷
۲۰۲	مدل طرح آزمایش‌ها	۱.۷
۲۰۵	برآزش α با تابع α_{ov}	۲.۲.۷
۲۰۷	مؤسسه‌های چندگانه	۳.۲.۷
۲۰۹	طرح آزمایش‌های عاملی	۳.۷
۲۱۶	مدل رگرسیونی با پاسخ‌های ردیاب	۴.۷
۲۲۳	روش‌های چندمتغیره‌ی پیوسته	۸
۲۲۳	استنباط آماری در مورد میانگین توزیع‌های چندمتغیره	۱.۸
۲۲۳	تولید بردارهای تصادفی از توزیع نرمال - چندمتغیره	۱.۱.۸
۲۲۵	آزمون نرمال چندمتغیره بودن مشاهده‌ها	۲.۱.۸
۲۲۶	محاسبه‌ی فاصله‌ی آماری در حالت چندمتغیره	۳.۱.۸
۲۲۷	آزمون فرض در مورد میانگین یک جامعه‌ی چندمتغیره	۴.۱.۸
۲۲۸	آزمون برابری میانگین‌های چند جامعه‌ی چندمتغیره	۵.۱.۸
۲۳۰	مؤلفه‌های اصلی	۲.۸
۲۳۳	تحلیل عاملی	۳.۸
۲۳۵	ضریب‌های همبستگی کانونی	۴.۸
۲۳۶	تحلیل ممیزی	۵.۸
۲۳۷	تحلیل ممیزی خطی	۱.۵.۸
۲۳۹	تحلیل ممیزی درجه‌ی دوم	۲.۵.۸
۲۴۰	خوشه‌بندی	۶.۸

۲۴۷	۹ سری‌های زمانی
۲۴۸	۱.۹ تحلیل و مدل‌سازی یک سری زمانی
۲۵۱	۲.۹ ایجاد یک سری زمانی در R
۲۵۳	۳.۹ رسم نمودار و بررسی اکتشافی یک سری زمانی
۲۶۰	۴.۹ تابع‌های خودهمبستگی و خودهمبستگی جزئی سری
۲۶۳	۵.۹ برازش مدل ARIMA
۲۷۳	کتاب‌ها
۲۷۵	واژه‌نامه
۲۷۹	نام‌نامه
۲۸۱	نمایه

پیش‌گفتار

در چند دهه‌ی اخیر پیاده‌سازی و اجرای روش‌های آماری برای تجزیه و تحلیل داده‌ها به کمک کامپیوتر در بیش‌تر علوم گسترش چشم‌گس داشته است. به همین دلیل نرم‌افزارهای آماری متنوعی ایجاد و مهارت کار با این نرم‌افزارها به یک مهارت ضروری برای پژوهشگران تبدیل شده است. توسعه‌ی روز افزون علم آمار با بهبود روش‌های آماری رایج و معرفی روش‌های جدید باعث شده است آماردانان به نرم‌افزارهایی گرایش یابند که علاوه بر اجرای محاسبات و تحلیل‌های آماری رایج، قابلیت پیاده‌سازی محاسبات و تحلیل‌های جدید را نیز داشته باشند.

در این میان، نرم‌افزار و زبان برنامه‌نویسی آماری R، به دلیل توانمندی و انعطاف‌پذیری بالا در پیاده‌سازی انواع محاسبات و تحلیل‌های آماری با استقبال زیادی روبه‌رو شده است. انعطاف‌پذیری برنامه‌نویسی در R به آماردانان و دیگر پژوهشگران کمک می‌کند تا محاسبات لازم بر روی داده‌ی خود را به سادگی پیاده‌سازی و ارزیابی کنند. به دلیل همین قابلیت‌ها، هر روز بسته‌های نرم‌افزاری جدید که توسط افراد گوناگون از سراسر دنیا با به‌کارگیری R توسعه داده شده‌اند، به شبکه‌ی جامع بایگانی R افزوده می‌شود و کم‌تر روش آماری را می‌توان یافت که در R قابل اجرا نباشد. این ویژگی‌ها و رایگان بودن R، سبب ریاضی دیگر نرم‌افزارهای آماری را که بیش‌تر تجاری هستند، از رقابت با R ناتوان کرده است. شاید به همین دلیل است که علاوه بر دانشگاه‌ها و مراکز پژوهشی سراسر جهان، سازمان‌های تجاری و صنعتی نیز به سمت به‌کار بستن R متمایل شده‌اند و نسخه‌های تجاری و اختصاصی نیز از R برای این طیف از کاربران توسعه داده شده است.

نیاز دانشجویان مقطع کارشناسی آمار و دانشجویان تحصیلات تکمیلی برخی رشته‌های دیگر به مرجعی به زبان فارسی در مورد R برای اجرای روش‌های آماری رایج و هم‌چنین نیاز دانشجویان تحصیلات تکمیلی آمار به یادگیری زبان برنامه‌نویسی R به منظور پیاده‌سازی کارهای پژوهشی خود، نگارندگان را به

تألیف این کتاب ترغیب کرد. بدیهی است که معرفی همه‌ی محاسبات و روش‌های آماری قابل اجرا در R بسیار فراتر از این کتاب و دانش نویسندگان آن است. این کتاب سعی بر آن دارد که با معرفی محاسبات و تحلیل‌های ابتدایی و ضروری قابل اجرا در R، زمینه‌ی آشنایی خواننده را با بخشی از قابلیت‌های R فراهم کند. پس از این آشنایی، خواننده می‌تواند با مراجعه به صفحه‌های راهنمای همراه R که به باور نگارندگان بهترین مرجع یادگیری R است، یادگیری خود را تکمیل کند. مرجع‌های ذکر شده در پایان کتاب نیز می‌توانند برای این منظور بسیار سودمند باشند.

این کتاب در بردارنده‌ی نه فصل است. فصل ۱ کتاب به معرفی نرم‌افزار و محاسبات مقدماتی در R می‌پردازد. ۲. خنارهای داده‌ای رایج در R در فصل ۲ و قابلیت‌های گرافیکی R در رسم انواع نمودارها در فصل ۳ معرفی شده‌اند. در فصل ۴ به نکته‌های ضروری برنامه‌نویسی در R به اختصار پرداخته شده است. در این فصل با تاکید بر تابع‌نویسی، مثال‌های متنوع و ساده‌ای از شیوه‌ی برنامه‌نویسی در R ارایه شده‌اند که خواننده می‌تواند با بررسی آن‌ها برنامه‌ها و تابع‌های جامع‌تری را بنویسد. اجرای روش‌های استنباط آماری کلاسیک برای یک رتبه‌بندی در R در فصل ۵ بحث شده است. اجرای تحلیل رگرسیونی خطی در R در فصل ۶ و تحلیل آماری در چند متغیر در فصل ۷ مورد اشاره قرار گرفته است. در فصل ۸ به کلیات اجرای روش‌های چندمتغیری، پوسته پرداخته و در فصل ۹ تحلیل ساده‌ی سری‌های زمانی به اختصار توضیح داده شده است.

شک نیست که کتاب حاضر خالی از نقص نیست. از همه‌ی خوانندگان و صاحب نظران گرامی صمیمانه خواهشمندیم با ارایه‌ی تذکرها، پیشنهادها و انتقاداتی که در دسترس ما قرار گیرد، کتاب در چاپ‌های بعدی یاری رسانند. در پایان لازم است از همه‌ی سازان و دوستانانی چون دکتر محی‌الدین ایزدی، دکتر مهرداد نیاپورست، دکتر حبیب جعفری، دکتر مریم شرفی، دکتر سیام‌الیزایی و آقای حسین ویسی‌پور که با ارایه‌ی نظرات خود ما را در ارتقای کیفی کتاب یاری رسانده‌اند، تشکر نماییم.

محمد مرادی و عبدالله جلیلیان

پاییز ۱۳۹۶