

محاسبات آماری با

S-PLUS

تألیف

انور محمدی

سرشناسه : محمدی، انور، ۱۳۶۱-
 عنوان و نام پدیدآور : محاسبات آماری با S-PLUS / تألیف انور محمدی.
 مشخصات نشر : تهران: گسترش علوم پایه، ۱۳۹۱.
 مشخصات ظاهری : ۳۷۲ ص.: مصور.
 شابک : ۹۷۸-۹۶۴-۴۹۰-۵۳۲-۲
 وضعیت فهرست : فیبا
 نویسی :
 موضوع : آمار -- برنامه‌های کامپیوتری
 موضوع : آمار ریاضی - نرم افزار
 شماره کتابشناسی : ۲۸۷۳۳۱۷
 ملی :

نام کتاب: محاسبات آماری با S-PLUS
 مؤلف: انور محمدی
 ناشر: انتشارات گسترش علوم پایه
 مدیر فنی: مهدی زنگنه
 طراحی جلد: حانمی کیا
 لیتوگرافی: توس
 چاپخانه: مهر
 سال نشر: ۱۳۹۱
 نوبت چاپ: اول
 شمارگان: ۱۰۰۰ جلد
 قیمت کتاب: ۹۷۸-۹۶۴-۴۹۰-۵۳۲-۲
 شابک: ISBN: 978-964-490-532-2

حق چاپ و نشر محفوظ و مخصوص ناشر می‌باشد.
 دفتر انتشارات و بخش شهرستان: میدان انقلاب، ابتدای کارگر جنوبی، کوچه مهدی‌زاده، پلاک ۹ طبقه همکف
 تلفن: ۱۵ ~ ۶۶۹۰۵۳۱۲ تلفکس: ۶۶۹۰۵۳۱۶-۰۲۱
 E-mail: Gostareh_op@yahoo.com
 www.Gostareh-pub.com

دفتر سفارشات تهران: خ آزادی، خ جمال‌زاده جنوبی، خ دیلمان، پلاک ۱۴ واحد ۲
 تلفن: ۱۴ ~ ۶۶۵۹۵۵۱۳
 فروشگاه فجر تهران: خ انقلاب، بین فروردین و اردیبهشت، پ ۱۴۶۲ - ۶۶۴۰۴۹۸۳ (۰۲۱)
 فروشگاه سخنکده: ضلع جنوب شرقی میدان انقلاب، بازار بزرگ کتاب طبقه زیرین پلاک ۲،
 (۰۲۱) ۶۶۴۵۳۰۰۲

www.ajansketab.com

(توجه) فروشنده و خواننده گرامی: این کتاب دارای برچسب اصالت کالا (مولوگرام) در روی جلد است که ضمن دقت در این مورد ما را در صورت عدم وجود مولوگرام مطلع سازید.
 با تشکر از همکاری شما

پیشگفتار

با گسترش نقش کامپیوتر در عرصه‌های محاسباتی، محققین علوم مختلف به توسعه نرم‌افزارهای مناسب برای استفاده از توانایی کامپیوترها در زمینه‌های مورد نیاز پرداختند. در علوم آماری نیز؛ نرم‌افزارهای گوناگونی برای پاسخگویی نیاز به محاسبات آماری پدید آمده‌اند. به علت اهمیت بالای محاسبات آماری و لزوم آشنایی فارغ‌التحصیلان رشته آمار با نرم‌افزارهای آماری که قابلیت انجام انواع محاسبات و روش‌های آماری را داشته باشند، درس «محاسبات آماری با کامپیوتر» در برنامه دوره کارشناسی این رشته گنجانده شده است. در حال حاضر و در میان نرم‌افزارهای آماری مناسب برای تدریس در این درس، بیشترین تمرکز معطوف به نرم‌افزار SPSS است. علت این امر اولاً ساده‌تر بودن کار با این نرم‌افزار در مقایسه با دیگر نرم‌افزارهای آماری برای دانشجویان دوره کارشناسی و ثانیاً قابلیت‌های متنوع آن در پوشش بسیاری از نیازهای آنهاست. بعلاوه به علت مشابهت ساختار برنامه‌نویسی آن با نرم‌افزار R، می‌تواند برای دانشجویان تحصیلات تکمیلی نیز آغاز خوبی باشد.

علت و انگیزه اصلی تالیف کتاب حاضر، فراهم نمودن امکان دسترسی دانشجویان و مدرسان محترم آماری به یک منبع مناسب و جامع فارسی برای پوشش مطالب مورد نیاز آنها در درس «محاسبات آماری با کامپیوتر» است. کتاب حاضر، حاصل مطالعات و علاقه مولف به مباحث محاسباتی رشته آمار و نیز تجربه چندین ترم تدریس در این زمینه است. بعلاوه در طی مراحل تالیف کتاب، نظرات جمعی از اساتید و همکاران محترم تکمیل‌کننده و ناظر مطالب تدوینی بوده است.

کتاب شامل ۱۱ فصل است که در ۳ دسته کلی قابل تقسیم می‌باشد. دسته اول شامل فصل‌های اول تا چهارم است که به آشنایی با S-PLUS و معرفی امکانات برنامه‌نویسی آن می‌پردازد. در این فصول سعی شده است با دسته‌بندی مناسب و منطقی مطالب و ارائه مثال‌های عملی، شرایط بگونه‌ای ایجاد گردد که دانشجویان تا حد معقولی توانایی نوشتن برنامه‌های مورد نیاز خود را در S-PLUS پیدا کنند.

دسته دوم شامل فصل‌های پنجم تا دهم کتاب است که در ابتدا نمودارهای آماری و سپس روش‌های آماری راجع را در فصول جداگانه بررسی می‌نماید. در این فصل‌ها علاوه بر اینکه مبانی مقدماتی روش‌ها و نحوه اجرای آنها در S-PLUS تشریح شده، سعی گردیده است با ارائه مثال‌های عملی و تمرینات مناسب، توانایی دانشجویان در استفاده از امکانات برنامه‌نویسی S-PLUS ارتقا یابد. مدرسین محترم می‌توانند به انتخاب خود بخشی از فصل‌های این دسته را حذف نموده یا در قالب پروژه‌های کلاسی در اختیار دانشجویان بگذارند. در پایان کتاب و در فصل یازدهم روش‌های مونت کارلو به صورت مختصری معرفی شده‌اند. مثال‌های مختلفی نیز از کاربردهای این روش‌ها ارائه گشته است.

در پایان بر خود لازم می‌دانم از زحمات و پشتیبانی‌های دوستان و همکارانی که در تکمیل این اثر مرا بسیار یاری نمودند، تشکر نمایم. از آقایان دکتر مهدی امیدی، دکتر شاهرز زارعی و دکتر امید کریمی که از آغاز کار نسخه‌های مختلف کتاب را مطالعه و نظرات اصلاحی بسیار مفیدی ارائه کردند، سپاس‌گذارم. از سایر همکارانم آقایان عیسی حسینی، مهدی شیرملکی، دکتر ربیع‌اله رحمانی، دکتر سیروس فتحی‌منش، دکتر نورالله تازی که می‌اندازه، دکتر افشین فلاح، دکتر امیر کاووسی، دکتر مهران نقی‌زاده، همکارانم در اداره مطالعات و تحقیقات اقتصادی بانک دی و سایر دوستانی که مشوق من در این راه بودند، کمال تشکر را دارم. همچنین از برخورد لازم می‌دانم از پشتیبانی‌های دلسوزانه مسئولان بانک دی و انتشارات گسترش علوم پایه تشکر نمایم.

بی‌شک اثر حاضر خالی از ایراد نیست. از خوانندگان کتاب تقاضا دارم نظرات و پیشنهادات خود را در این زمینه از طریق آدرس a.m.stat@gmail.com با بنده در میان بگذارند.

انور محمدی

مرداد ۱۳۹۱

فهرست مطالب

۱۵	فصل اول: آشنایی مقدماتی با S-PLUS
۱۵	۱-۱ مقدمه
۱۶	۲-۱ تاریخچه S و S-PLUS
۱۷	۳-۱ شروع کار با S-PLUS
۲۰	۴-۱ چند مفهوم اولیه در برنامه‌نویسی
۲۳	فصل دوم: ساختار داده‌ها در S-PLUS
۲۳	۱-۲ نحوه نامگذاری و تعریف ساختار داده
۲۴	۲-۲ بردار
۲۵	۱-۲-۲ ساخت بردارهای ساده
۲۶	۲-۲-۲ ویژگی‌های یک بردار
۲۷	۳-۲-۲ ساخت بردارهای الگودار
۳۰	۴-۲-۲ ساخت بردارهای رشته‌ای
۳۱	۵-۲-۲ اشاره به محتویات بردار
۳۲	۳-۲ ماتریس
۳۳	۱-۳-۲ تابع <code>matrix()</code>
۳۵	۲-۳-۲ پیوند ماتریس‌ها
۳۵	۳-۳-۲ سایر ابزارهای ساخت ماتریس
۳۶	۴-۳-۲ اشاره در ماتریس

۳۷	۴-۲ آرایه
۳۸	۵-۲ لیست
۳۹	۱-۵-۲ استفاده از لیست‌ها جهت اشاره در ماتریس
۴۰	۶-۲ چارچوب داده
۴۱	۱-۶-۲ تابع attach()
۴۱	۷-۲ سایر ساختار داده‌ها
۴۲	۱-۷-۲ فاکتور
۴۲	۲-۷-۲ رسته
۴۳	۳-۷-۲ سری زمانی
۴۴	۸-۲ از نمایی و تبدیل نوع و ویژگی‌های ساختار داده‌ها
۴۵	۹-۲ وارد کردن داده‌ها در S-PLUS
۴۵	۱-۹-۲ وارد کردن در دستورات ساخت ساختار داده
۴۵	۲-۹-۲ وارد کردن در صفحه داده
۴۷	۳-۹-۲ وارد کردن با تابع scan()
۴۸	۴-۹-۲ خواندن داده‌ها از یک فایل متنی
۵۳	فصل سوم: محاسبات ریاضی
۵۳	۱-۳ عملگرهای ریاضی
۵۴	۱-۱-۳ کاربرد عملگرهای ریاضی برای بردارها
۵۵	۲-۱-۳ اولویت محاسباتی عملگرهای ریاضی
۵۵	۲-۳ توابع ریاضی
۵۵	۱-۲-۳ توابع نمایی
۵۶	۲-۲-۳ توابع جزء صحیح
۵۶	۳-۲-۳ توابع مثلثاتی
۵۷	۴-۲-۳ توابع فاکتوریل
۵۷	۵-۲-۳ توابع کار با اعداد مختلط
۵۸	۶-۲-۳ توابع مرتب‌سازی و رتبه
۶۰	۷-۲-۳ توابع مشتق‌گیری و انتگرال‌گیری
۶۱	۸-۲-۳ سایر توابع
۶۱	۳-۳ محاسبات ماتریسی

۶۳	۱-۳-۳ ضرب و عملیات خارجی روی ماتریس‌ها
۶۳	۲-۳-۳ معکوس کردن و تجزیه ماتریس‌ها
۶۵	۴-۳ توابع آماری و احتمالی
۶۵	۱-۴-۳ توابع احتمالی
۶۶	۲-۴-۳ نگرانی در تولید مقادیر تصادفی
۶۷	۳-۴-۳ شاخص‌های آماری
۶۹	۴-۴-۳ شاخص‌های آماری برای ماتریس‌ها
۶۹	۵-۴-۳ تابع <code>apply()</code>
۷۳	فصل چهارم: اصول برنامه‌نویسی در زبان S
۷۳	۱-۴ دستورات شرطی
۷۴	۱-۱-۴ عبارت منطقی
۷۵	۲-۱-۴ دستور <code>if</code>
۷۷	۳-۱-۴ دستور <code>if-else</code>
۷۸	۴-۱-۴ بلوک دستورات
۷۹	۲-۴ دستورات تکراری (حلقه‌ها)
۷۹	۱-۲-۴ حلقه <code>while</code>
۸۱	۲-۲-۴ حلقه <code>repeat</code>
۸۳	۳-۲-۴ حلقه <code>for</code>
۸۴	۴-۲-۴ کنترل در درون حلقه
۸۴	۳-۴ ساخت تابع جدید
۸۷	۱-۳-۴ ساختن تابع به صورت عملکرد
۹۱	فصل پنجم - نمودارهای آماری
۹۱	۱-۵ تابع <code>plot()</code>
۹۲	۱-۱-۵ شناسه <code>type</code>
۹۵	۲-۱-۵ تعیین نوع نقاط و خطوط
۹۷	۳-۱-۵ تعیین عناوین و برچسب‌های نمودار
۹۹	۴-۱-۵ تعیین حدود محورها
۹۹	۵-۱-۵ افزودن متن به یک نمودار
۱۰۰	۲-۵ اضافه کردن منحنی به نمودار موجود

۱۰۱	۳-۵ بافت‌نگار (هستوگرام)
۱۰۳	۱-۳-۵ خروجی تابع <code>hist()</code> غیر از نمودار بافت‌نگار
۱۰۴	۲-۳-۵ رسم منحنی روی نمودار بافت‌نگار
۱۰۵	۳-۳-۵ منحنی چگالی تجربی
۱۰۵	۴-۵ نمودار جعبه‌ای
۱۰۶	۱-۴-۵ نمودار جعبه‌ای چندگانه
۱۰۶	۲-۴-۵ خروجی تابع <code>boxplot()</code> غیر از نمودار جعبه‌ای
۱۰۷	۵-۵ نمودارهای میله‌ای و دایره‌ای
۱۰۷	۱-۵-۵ جدول فراوانی
۱۰۸	۲-۵-۵ نمودار میله‌ای
۱۰۸	۳-۵-۵ نمودار دایره‌ای
۱۱۱	۶-۵ نمودار چارک-چارک
۱۱۱	۱-۶-۵ مقایسه توزیع دوسری داده
۱۱۲	۲-۶-۵ ارزیابی توزیع یک سری داده
۱۱۳	۳-۶-۵ ارزیابی نرمال بودن توزیع یک سری داده
۱۱۴	۷-۵ نمودارهای ۳ بعدی
۱۱۴	۱-۷-۵ نمودار چشم‌انداز ۳-بعدی
۱۱۵	۲-۷-۵ نمودارهای تخت برای نمایش داده‌های ۳-بعدی
۱۲۱	فصل ششم: آزمون‌های آماری، مقایسه میانگین‌ها
۱۲۲	۱-۶ بررسی نرمال بودن
۱۲۲	۱-۱-۶ نمودارهای بافت‌نگار و جعبه‌ای
۱۲۳	۲-۱-۶ نمودار $Q-Q$ (چارک-چارک)
۱۲۴	۳-۱-۶ آزمون‌های نرمال بودن
۱۲۵	۲-۶ آزمون‌های تک نمونه‌ای
۱۲۶	۱-۲-۶ آزمون t تک نمونه‌ای با تابع <code>t.test()</code>
۱۲۸	۲-۲-۶ خروجی تابع <code>t.test()</code>
۱۲۹	۳-۲-۶ آزمون ویلکاکسون تک نمونه‌ای
۱۳۱	۳-۶ آزمون‌های دو نمونه‌ای
۱۳۲	۱-۳-۶ بررسی پیش فرض‌ها

۱۳۲	۲-۳-۶ آزمون دو نمونه‌ای با تابع t test
۱۳۵	۳-۳-۶ نمونه‌های جفتی
۱۳۷	۴-۳-۶ آزمون ویلکا کسون برای مقایسه دو نمونه مستقل و زوجی
۱۳۷	۴-۶ آزمون‌های چند نمونه‌ای
۱۴۱	فصل هفتم: مدل‌های رگرسیونی
۱۴۲	۱-۷ ضرایب همبستگی
۱۴۲	۱-۷ ضریب همبستگی خطی
۱۴۲	۲-۱-۷ نمودار پراکنش
۱۴۳	۳-۱-۷ برآورد ضریب همبستگی پیرسون
۱۴۴	۴-۱-۷ ضریب همبستگی رتبه‌ای اسپیرمن
۱۴۵	۵-۱-۷ ضریب تانگن
۱۴۵	۲-۷ رگرسیون خطی چندگانه
۱۴۶	۱-۲-۷ بیش‌فرض‌ها
۱۴۷	۲-۲-۷ برآورد پارامترهای مدل
۱۴۷	۳-۲-۷ خروجی‌های تابع $lm()$
۱۵۱	۴-۲-۷ بررسی مانده‌ها
۱۵۷	۳-۷ انتخاب متغیرها در مدل رگرسیونی
۱۵۸	۱-۳-۷ تابع $leaps()$
۱۶۱	۲-۳-۷ تابع $stepwise()$
۱۶۴	۴-۷ رگرسیون استوار
۱۶۴	۱-۴-۷ رگرسیون کمترین توان‌های دوم پیراسته
۱۶۵	۲-۴-۷ رگرسیون کمترین میانه توان‌های دوم
۱۶۶	۳-۴-۷ رگرسیون کمترین قدر مطلق مانده‌ها
۱۶۶	۴-۴-۷ ام-برآوردگر رگرسیون
۱۶۷	۵-۴-۷ مقایسه روش‌های رگرسیون استوار
۱۶۸	۶-۴-۷ ام-ام برآوردگر رگرسیون
۱۶۸	۵-۷ سایر مدل‌های رگرسیونی
۱۶۸	۱-۵-۷ رگرسیون کمترین توان‌های دوم موزون
۱۶۹	۲-۵-۷ رگرسیون غیرخطی

فصل هشتم: طرح و تحلیل آزمایش‌ها

- ۱۷۵ ۱-۸ آنالیز واریانس یک طرفه
- ۱۷۷ ۱-۱-۸ بررسی پیش فرض‌ها
- ۱۷۸ ۲-۱-۸ بررسی نموداری
- ۱۷۸ ۳-۱-۸ تابع $\text{aov}()$
- ۱۸۲ ۴-۱-۸ مقایسات چندگانه
- ۱۸۳ ۵-۱-۸ آزمون کروסקال-والیس
- ۱۸۴ ۲-۸ آزمایش تک عاملی
- ۱۸۴ ۱-۲-۸ اثرهای ثابت، اثرهای تصادفی
- ۱۸۵ ۳-۸ آزمایش دو عاملی
- ۱۸۵ ۱-۳-۸ آماده سازی داده‌ها
- ۱۸۶ ۲-۳-۸ بررسی نموداری داده‌ها
- ۱۸۶ ۳-۳-۸ مدل آماری و آنالیز واریانس
- ۱۸۷ ۴-۳-۸ فرمول مدل
- ۱۸۷ ۵-۳-۸ مثال عملی
- ۱۹۲ ۴-۸ طرح عاملی 2^k

فصل نهم: روش‌های چندمتغیره

- ۱۹۹ ۱-۹ تحلیل مولفه‌های اصلی
- ۲۰۰ ۱-۱-۹ برآورد ماتریس کوواریانس
- ۲۰۲ ۲-۱-۹ تابع $\text{princomp}()$
- ۲۰۳ ۳-۱-۹ مثال عملی
- ۲۰۴ ۴-۱-۹ نمودارها در تحلیل مولفه‌های اصلی
- ۲۰۶ ۲-۹ تحلیل عاملی اکتشافی
- ۲۰۶ ۱-۲-۹ برآورد بارهای عاملی
- ۲۰۷ ۲-۲-۹ تابع $\text{factanal}()$
- ۲۰۸ ۳-۲-۹ دوران عاملی
- ۲۰۹ ۴-۲-۹ امتیازهای عاملی
- ۲۰۹ ۵-۲-۹ مثال عملی
- ۲۱۰ ۳-۹ تحلیل خوشه‌ای

۲۱۱	۹-۳-۱ فاصله
۲۱۲	۹-۳-۲ خوشه‌بندی سلسله‌مراتبی
۲۱۶	۹-۳-۳ خوشه‌بندی غیر سلسله‌مراتبی
۲۲۳	فصل دهم: تحلیل داده‌های گسته
۲۲۳	۱۰-۱ آزمون‌های یک متغیری
۲۲۳	۱۰-۱-۱ آزمون دقیق دو جمله‌ای
۲۲۴	۱۰-۱-۲ آزمون نسبت
۲۲۶	۱۰-۲ جدول توافقی
۲۲۸	۱۰-۳ آزمون‌های استقلال
۲۲۸	۱۰-۳-۱ آزمون خی دو پیرسون
۲۲۹	۱۰-۳-۲ آزمون دقیق فشر
۲۳۰	۱۰-۳-۳ آزمون مک‌نمار
۲۳۳	۱۰-۳-۴ آزمون کوکران-منتال-مانزِل
۲۳۹	فصل یازدهم: روش‌های مونت کارلو
۲۴۰	۱۱-۱ تولید اعداد تصادفی
۲۴۰	۱۱-۱-۱ اعداد شبه تصادفی
۲۴۰	۱۱-۱-۲ روش هم‌نهشی ضربی
۲۴۲	۱۱-۱-۳ مقادیر تصادفی در S-PLUS
۲۴۲	۱۱-۲ تولید مقدار تصادفی از توزیع دلخواه
۲۴۲	۱۱-۲-۱ روش تبدیل وارون
۲۴۳	۱۱-۲-۲ روش تبدیل وارون برای متغیرهای تصادفی گسته
۲۴۵	۱۱-۲-۳ روش نمونه‌گیری رد و پذیرش
۲۴۷	۱۱-۲-۴ روش قطبی برای تولید مقدار از توزیع نرمال استاندارد
۲۴۸	۱۱-۳ چند کاربرد روش‌های مونت کارلو
۲۴۸	۱۱-۳-۱ محاسبه مقدار تقریبی انتگرال‌ها
۲۵۰	۱۱-۳-۲ شبیه‌سازی مسائل احتمالی
۲۵۵	پیوست ۱: طریقه نصب S-PLUS 8.0
۲۶۳	پیوست ۲: روش استفاده از S-PLUS Language Reference