

استفاده از R در

آمار مقدماتی

نویسنده: برای تمام رشته‌های

علوم پایه، مهندسی

کارشناسی، کارشناسی ارشد، دکترا

مؤلف:

JOHN VERZAN

مترجمین:

صغری طاهرخانی

نجم‌الدین گنج‌خانلو - علی فصیحی

سرشناسه	ورزنی جان
عنوان و نام پدیدآور	Verzani, John
مشخصات نشر	استفاده از R در آمار مقدماتی قابل استفاده برای رشته‌های علوم پایه، مهندسی... مولف ؟
مشخصات ظاهری	مترجمان صغری طاهرخانی، نجم‌الدین گنج‌خانلو، علی فصیحی
شابک	کرج: آوای ویانا، ۱۳۹۳.
وضعیت فهرست‌نویسی	۲۰۸ص: جدول، نمودار.
داشت	978-600-93954-9-1
موضوع	فیبا
موضوع	عنوان اصلی: Using R for introductory statistics, 2014
موضوع	آمار - داده‌پردازی
موضوع	آر (زبان برنامه‌نویسی کامپیوتر)
موضوع	طاهرخانی، صغری، ۱۳۶۱ - مترجم
موضوع	گنج‌خانلو، نجم‌الدین، ۱۳۵۷ - مترجم
موضوع	فصیحی، علی، ۱۳۵۷ - مترجم
رده‌بندی کنگره	QA۲۷۶/۴۵/۱۴
رده‌بندی دیویی	۵۱۹/۵۰۲/۰۵۳۳
شماره کتابشناسی ملی	۳۸۰۰۱



نام کتاب	R در آمار مقدماتی
مولف	JOHN VERZANI
مترجمین	صغری طاهرخانی - نجم‌الدین گنج‌خانلو - علی فصیحی
ناشر	انتشارات آوای ویانا
نوبت چاپ	اول - ۱۳۹۳
شمارگان	۱۰۰۰ نسخه
طراح جلد	حسین فصیحی
قیمت	۸۰۰۰ تومان
شابک	۹۷۸-۶۰۰-۹۳۹۵۴-۹-۱

دفتر انتشارات: کرج - جهانشهر - خیابان استانداری البرز - تلفن تماس ۴۴۲۱۵۸۹

تلفن همراه: ۰۹۱۲۲۶۰۷۶۷۱ ۰۹۳۳۷۶۹۷۹۶۴ ۰۹۲۱۳۱۳۵۲۸۶

وب سایت: www.avaiveyana.com پست الکترونیک: avaiveyana@gmail.com

info@avaiveyana.com

هرگونه نسخه‌برداری و تکثیر از کل کتاب یا بخشی از آن ممنوع است و پیگرد قانونی دارد.

مقدمه

این کتاب حاصل کار عملی و تجربه سالهای ما است در این کتاب سعی کرده‌ایم برای دانشجوی آماری زمینه‌ای فراهم سازیم تا در ابتدای مسیر و ورود به این رشته، تمام آموخته‌های خود را با نرم‌افزار R انجام دهد و کیفیت و کمیت کار خود را با سایر نرم‌افزارهای موجود در آمار از جمله SPSS, MINETAP, SAS, S PLOT مقایسه کند و در برنامه نوشتن پروژه‌های مقاطع بالاتر خصوصاً اقدام نماید از آنجا که این کار یک کار ترجمه از کتاب لاتین و منبع انگلیسی می‌باشد لذا ما تمام مثالها را بدون تغییر در متن آورده‌ایم تا دانشجو با داشتن هر دو منبع بتواند در بالا بردن ارتقای آموخته‌های خود قدم بردارد لذا این کتاب را در پانزده فصل تنظیم کرده‌ایم امید است با مطالعه و تسلط بر این کتاب بتوانیم در ارتقای سطح علمی شما سروران بزرگوار قدمی پر بار و مفید فراهم سازیم.

التماس دعا

صغری طاهرخانی - نجم الدین گنج خانی - علی فصیحی

پیشگفتار

در این کتاب روش استفاده از بسته نرم افزار آماری R در رشته آمار مقدماتی مورد بررسی قرار می گیرد، هدف اصلی ما استفاده از یک کتاب راهنما برای نشان دادن ویژگیهای R است که بتوان در یک نیمسال رشته آمار مقدماتی آن را یاد گرفت و این مطالب برای بهره وری از ورژن R ۱/۵/۰ یا پس از آن نوشته شده است.

در ابتدا نکته را متذکر می شویم؛ چندین نوع مرجع داده ها و عملیات وجود دارد که قبل از استفاده از R به نصب و راه اندازی دارند. نصب این داده ها راحت است، اما دستورالعمل ها بسته به سیستم خواس و سخت افزار است. مثلاً برای کاربران ویندوز شما نیاز دارید تا فایل زیپ را دانلود کنید و سپس از منوی «packages» آن را نصب کنید.

در اینجا دو آدرس جهت کمک به دانلود فایل های، مجموعه داده، اختصاص داده ایم

<http://www.math.csi.cuny.edu/Statistics/R/simpleR/SimpleR.zip> for Windows users.

<http://www.math.csi.cuny.edu/Statistics/R/simpleR/SimpleR.4.tar.gz> for UNIX users.

اگر لازم باشد فایل را می توان به یک آدرس اینترنتی فرستاد. علاوه تک تک مجموعه داده ها را می توان بصورت اینترنتی در دفترچه راهنما پیدا کرد.

<http://www.math.csi.cuny.edu/Statistics/R/simpleR/SimpleR>.

توجه کنید باید جدیدترین ورژن را از آدرس زیر بررسی کنید:

<http://www.math.csi.cuny.edu/Statistics/R/simpleR>

۱- در یونیکس یک استفاده از فرمان `R CMD TN STALL packagename.tar.gz` برای برخی از مجموعه داده ها

از نویسندگان دیگر مثل کیچنز فرض گرفته شده است

فهرست مطالب

صفحه

عنوان

۱۵	فصل اول / R چیست؟
۱۶	تفاوت با نرم افزارهای دیگر، کمبودهای R چیست؟
۱۶	داده
۱۶	را سازی R
۱۷	for و while در R
۱۷	وارد کردن داده ها
۱۸	استعمال یک دستور
۱۸	داده یک بردار است
۲۰	آیا می توانیم R را در یک سوه تمیز انجام دهیم
۲۰	چگونه می توانیم شاخص ها (صفحات) تنای با مقادیر درست را بیابیم؟
۲۲	چگونه یک ذخیره را ثبت کنید، به آن اضافه کنید
۲۲	اساس های R: میانجی ثبت داده های گرافیک
	حداکثر مقدار ذخیره چیست؟
۲۵	مسائل
۲۷	فصل دوم / داده های یک متغیری
۲۷	داده های مطلق:
۲۸	استفاده از جدول ها
۲۹	فاکتورها
۲۹	نمودار میله ای
۳۱	نمودار کلوچه ای (دایره ای)
۳۳	داده های عددی
۳۳	سنجش عددی مرکز و پراکندگی
۳۴	تفاوت بین مقادیر عددی و مقادیر کمی
۳۵	سنجش های پایدار مرکز و پراکندگی
۳۶	نمودار شاخه و برگ

۳۷	 بیسیک های R: apropos , help
۳۷	 ساختن داده های عددی مطلق
۳۸	 پیشینه نما (هیستوگرام)
۴۰	 نمودار جعبه ای
۴۱	 بیسیک های R:
۴۳	 نمودار چند ضلعی فراوانی
۴۴	 چگالی ها
۴۵	 مسائل
۴۷	 فصل سوم: داده های دو متغیری
۴۷	 داده های مطلق دو متغیری
۴۹	 رسم داده های جدول
۵۰	 نسبت های مشروط
۵۰	 کار کردن با داده های دو متغیر: مابین های مطلق در مقابل متغیرهای عددی
۵۲	 داده های دو متغیری: عددی در مقابل عددی
۵۲	 مقایسه دو توزیع با نمودارها
۵۲	 استفاده از جدول نواری یا نمودار نقطه ای
۵۳	 استفاده از نمودار پراکندگی برای مقایسه روابط
۵۴	 بیسیک های R: الصاق کردن چه کاری انجام می دهد؟
۵۵	 برگشت خطی
۵۶	 دستور abline
۵۷	 نمودارهای رسوبی (باقیمانده)
۵۸	 ضریب همبستگی
۵۹	 تعیین محل نقاط
۶۲	 برگشت پایدار
۶۲	 استفاده از rlm یا las برای برگشت پایدار
۶۴	 بیسیک های R: طراحی نمودارها با استفاده از R
۶۴	 اضافه کردن به یک نمودار با curve و lines و points
۶۵	 مسائل

فصل چهارم / داده‌های چند متغیری ۶۹

ذخیره داده‌های چند متغیری در چارچوب داده ها ۶۹

دستیابی به داده‌ها در چارچوب داده ای ۷۰

دسترسی به یک آرایش ۷۱

دستیابی به یک فهرست ۷۲

ستکاری در چارچوب داده ها: stack , unstack ۷۲

ماه از سانه‌گذاری فرمولی مدل R ۷۳

زمانی که دو متغیر است response ~ predictor ۷۴

رسم‌های بررسی داده‌های چند متغیری ۷۴

جدول‌های استکی روشی ۷۴

جدول‌های آری ۷۶

نمودارهای ویونی و نوارهای چگالی ۷۷

نمودارهای پراکنده ۷۸

نمودار پراکندگی جفتی ۸۰

بسته بشک‌ای ۸۱

مسائل ۸۳

فصل پنجم / داده‌های تصادفی ۸۵

تولیدکننده‌های عدد تصادفی در R - تابع r ۸۵

اعداد یکنواخت ۸۶

اعداد طبیعی ۸۷

توزیع دو جمله ایی اعداد تصادفی ۸۸

توزیع نمایی ۸۹

مدلسازی با جایگزینی و بدون جایگزینی با استفاده از sample ۹۰

یک نمونه خود راه انداز ۹۱

توابع q , p , d ۹۲

استاندارد سازی، مقیاس و اعداد Z ۹۳

مسائل ۹۴

فصل ششم / شبیه‌سازی ۹۷

قضیه حد مرکزی ۹۷

۹۹	حلقه‌های لوپ
۹۹	بسیک‌های R: دستورالعمل برای for
۱۰۱	ترسیم‌های نرمال
۱۰۱	استفاده از توابع و simile.sim
۱۰۵	مسائل
۱۰۹	فصل نهم / تحلیل داده‌های اکتشافی
۱۰۹	جعبه اینشتین
۱۰۹	نمودارهای برای داده‌های مطلق
۱۱۶	مسائل
۱۱۷	فصل هشتم / تخمین فاصله اعتماد
۱۱۷	نظریه نسبت جمعیتی
۱۲۱	آزمایش نسبت
۱۲۱	آزمون prop دقیق‌تر است.
۱۲۲	آزمایش Z
۱۲۳	آزمایش t
۱۲۴	مقایسه مقادیر p از z, t
۱۲۵	فاصله اطمینان برای میانگین
۱۲۶	مسائل
۱۲۹	فصل نهم / آزمون فرضیه‌ها
۱۲۹	آزمایش یک پارامتر جمعیت
۱۳۱	آزمون‌های میانگین
۱۳۱	یک آزمون مناسب برای مرکز چیست؟
۱۳۲	آزمون‌های رتبه‌ای
۱۳۳	مسائل
۱۳۵	فصل دهم / آزمون‌های دو نمونه‌ای
۱۳۶	آزمون‌های دو نمونه‌ای t
۱۳۶	واریانس‌های برابر
۱۳۷	واریانس‌های نابرابر
۱۳۸	نمونه‌های هم‌ت

آزمون‌های دو نمونه‌ای مقاوم ۱۳۹

مسائل ۱۴۰

فصل یازدهم / آزمون «خی دو» ۱۴۳

توزیع «خی دو» ۱۴۳

چرا X^2 ؟ چه چیز سبب می‌شود که در آمار توزیع X^2 داشته باشد؟ ۱۴۶

آزمون‌های داده‌های مستقل «خی دو» ۱۴۷

آزمون‌های t برای همگونی ۱۴۷

مسائل ۱۴۹

فصل دوازدهم / تحلیل رگرسیون ۱۵۳

مدل رگرسیون خطی ساده ۱۵۳

آزمایش فرضیه مدل ۱۵۶

باقیمانده‌ها در مقایسه خطی ۱۵۷

استنتاج آماری ۱۵۷

درباره σ^2 ۱۵۷

استنتاج درباره B_1 ۱۵۸

استنتاج درباره B_0 ۱۵۹

فاصله‌های اعتماد ۱۶۰

دستورهای سطح پایین R ۱۶۰

یافتن باند‌های اعتماد ۱۶۳

مسائل ۱۶۴

فصل سیزدهم / رگرسیون خطی چندگانه ۱۶۵

مدل ۱۶۶

رسم اسانتز ۱۶۶

مسائل ۱۷۴

فصل چهاردهم / تحلیل واریانس ۱۷۵

تحلیل یک طرفه واریانس ۱۷۵

تحلیل واریانس توصیف شده ۱۷۷

میانگین مجموع مربعات ۱۷۹

استفاده از av ۱۸۰

۱۸۰	آزمون کروسکال هوالیس
۱۸۰	مسائل:
۱۸۳	فصل پانزدهم / پیوست نصب R
۱۸۳	وب سایت اصلی R
۱۸۳	پیوست: بسته‌های خارجی
۱۸۴	پیوست: یک نمونه R
۱۸۴	یک نمونه شامل رگرسیون
۱۸۵	لود کردن مجموعه داده‌های mtcars
۱۸۵	دستیابی به داده‌ها
۱۸۵	داده‌های مطلق
۱۸۶	داده‌های عددی
۱۸۷	تغییرپذیری یا پراکندگی
۱۸۷	روابط دو متغیری
۱۸۹	آزمایش فرضیات برگشت
۱۹۰	مرتب کردن
۱۹۰	آزمون‌های t
۱۹۰	اتصال مجموعه داده‌ای
۱۹۰	EDA
۱۹۲	محصور کردن
۱۹۲	یک مثال شبیه سازی شده
۱۹۴	پیوست: وقتی R راه اندازی می‌شود چه اتفاقی می‌افتد؟
۱۹۴	پیوست: استفاده از توابع
۱۹۵	الگوی اصلی
۱۹۵	اسم تابع
۱۹۵	تابع کلید
۱۹۵	آرگومان‌های تابع
۱۹۵	بدون آرگومان
۱۹۶	یک آرگومان پیش فرض
۱۹۸	بدنه تابع و مقدار برگشت تابع
۱۹۸	حلقه‌های loop
۱۹۹	توضیحات مشروط

۲۰۰	پیوست: وارد کردن داده‌ها به R
۲۰۰	استفاده از C
۲۰۰	استفاده از اسکن
۲۰۱	استفاده از scan یک فایل
۲۰۱	ویرایش داده‌های شما
۲۰۳	فایل‌های دارای پهنای ثابت
۲۰۳	فایل‌های صفحه گسترده
۲۰۴	پیوست: ترفندهای تدریس