

با نام و یاد او

تحلیل رگرسیون خطی با R

تألیف

سمانه افتخاری مهابادی

استادیار آمار دانشگاه تهران

سید مرتضی امینی

استادیار آمار دانشگاه تهران

احمد پارسیان

استاد بازنشسته آمار دانشگاه تهران

حسین زمانی

استادیار آمار دانشگاه هرمزگان

انتشارات علمی پارسیان

عنوان و نام پدیدآور
 مشخصات نشر
 مشخصات ظاهری
 شابک
 وضعیت فهرست نویسی
 یادداشت
 یادداشت
 موضوع
 موضوع
 موضوع
 موضوع
 شناسه افزودن
 رده‌بندی کنگره
 رده‌بندی دی‌یو
 شماره کتابشناسی ملی

تحلیل رگرسیون خطی با R / تألیف سمانه افتخاری مهابادی ... [و دیگران].
 تهران: انتشارات علمی پارسیان، ۱۳۹۷.
 ۲۹۱ ص.
 ۲ - ۵۶ - ۹۴۴۸ - ۶۰۰ - ۹۷۸
 فیبا
 تألیف سمانه افتخاری مهابادی، سید مرتضی امینی، احمد پارسیان، حسین زمانی.
 کتابنامه.
 برگشت (ریاضیات) - - برنامه‌های کامپیوتری
 Regression analysis - - Computer programs
 آر (زبان برنامه‌نویسی کامپیوتر)
 R (Computer program language)
 افتخاری مهابادی، سمانه، ۱۳۶۴ -
 QA۲۷۸/۲/۳ ۱۳۹۷
 ۵۱۹/۵۳۶
 ۵۱۴۱۲۶۲

ادشارات علمی پارسیان

بهران: صندوق پستی ۳۷۹۵ - ۱۶۷۶۵ (مدیر داخلی و تولید: مهدیه شهبازی)
 تلفن ثابت: ۸۸۰۶۸۳۳۴
 سفن همراه: ۰۹۱۲۷۰۱۸۲۹ و ۰۹۳۶۶۰۳۳۱۷۰
 پیام‌رسان: manager@parsiansp.ir



انتشارات علمی پارسیان
 ۱۰۵

تحلیل رگرسیون خطی با R	نام کتاب
سمانه افتخاری مهابادی، سید مرتضی امینی، احمد پارسیان و حسین زمانی	مؤلفان
مهدیه شهبازی	ویراستار
سید مرتضی امینی	حروف‌نگار
سمانه افتخاری مهابادی	صفحه‌آرا
عصر مجلسی‌راد	طراح جلد
انتشارات علمی پارسیان	ناشر
اول، بهار ۱۳۹۷	نوبت چاپ
جلد ۶۰۰	شمارگان
۲ - ۶ - ۹۴۴۸۵ - ۶۰۰ - ۹۷۸	شابک
۳۴,۵۰۰ تومان	قیمت

حق چاپ برای انتشارات علمی پارسیان محفوظ است.

فهرست مطالب

پیش‌گفتار	ث
۱ رگرسیون خطی با n متغیر پیشگو	۱
۱.۱ رابطه‌ی بین متغیرهای X و Y	۲
۲.۱ تحلیل رگرسیونی	۳
۱.۲.۱ نمودار پراکنش	۴
۳.۱ مدل‌های رگرسیونی	۱۲
۴.۱ برآورد β_0 و β_1 در مدل رگرسیون خطی ساده	۱۸
۵.۱ ویژگی‌های برآوردگرهای کم‌ترین توان دوم β_0 و β_1	۲۳
۱.۵.۱ برآورد نقطه‌ای میانگین شرطی متغیر پاسخ	۲۶
۶.۱ مانده‌ها	۳۰
۷.۱ مجموع توان دوم خطا	۳۴
۸.۱ برآورد واریانس جمله‌ی خطا	۳۵
۹.۱ برآورد ماکسیمم درستنمایی پارامترهای مدل رگرسیون خطی ساده	۳۸
۱۰.۱ رگرسیون گذرا از مبدأ (بدون عرض از مبدأ)	۴۳
۱۱.۱ مدل رگرسیون خطی برای متغیر پیشگوی کیفی دوسطحی	۴۵
۱۲.۱ تمرین‌ها	۴۷
۲ استنباط در مدل رگرسیون خطی ساده	۵۳
۱.۲ استنباط درباره‌ی پارامترهای مدل رگرسیون خطی ساده	۵۳
۱.۱.۲ استنباط درباره‌ی شیب خط رگرسیونی	۵۳
۲.۱.۲ استنباط درباره‌ی عرض از مبدأ خط رگرسیونی	۶۲
۲.۲ استنباط درباره‌ی شیب خط رگرسیونی گذرا از مبدأ	۶۹
۳.۲ بازه‌ی اطمینان برای میانگین شرطی متغیر پاسخ	۷۰

۴.۲	بازوی پیش‌بینی برای یک مقدار جدید متغیر پاسخ	۷۵
۵.۲	استنباط درباره‌ی واریانس خطاها	۷۹
۶.۲	تحلیل واریانس	۸۰
۷.۲	آزمون خطی کلی	۸۵
۸.۲	تحلیل همبستگی	۸۷
۹.۲	قضیه‌ی گاوس-مارکف	۹۴
۱۰.۲	تمرین‌ها	۹۷
۳	روش‌های تشخیصی و عیب‌یابی مدل رگرسیون خطی ساده	۱۰۳
۱.۳	بررسی درستی فرض‌های مدل رگرسیون خطی	۱۰۳
۲.۳	بررسی نرمال بودن مانده‌ها	۱۰۴
۱.۲.۳	بررسی نموداری فرض خطی بودن تابع رگرسیونی	۱۰۴
۲.۲.۳	بررسی همبستگی واریانس‌ها	۱۱۱
۳.۲.۳	بررسی فرض ناهمبسته بودن خطاها	۱۱۷
۴.۲.۳	بررسی نموداری فرض نرمال بودن خطاها	۱۲۰
۳.۳	داده‌های دورافتاده	۱۲۴
۴.۳	آزمون‌های فرض بر اساس مانده‌ها	۱۲۷
۱.۴.۳	آزمون نقص برازش مدل خطی (آزمون فرض خطی بودن)	۱۲۷
۲.۴.۳	آزمون فرض همبستگی واریانس خطاها	۱۳۳
۳.۴.۳	آزمون فرض ناهمبسته بودن خطاها	۱۳۴
۴.۴.۳	آزمون فرض نرمال بودن خطاها	۱۳۶
۵.۳	معیارهای توصیفی	۱۳۷
۶.۳	مدل‌های رگرسیونی خطی ذاتی	۱۴۰
۷.۳	تبدیل‌های ثابت‌سازی واریانس خطاها	۱۴۸
۱.۷.۳	روش کم‌ترین توان دوم وزنی	۱۵۰
۸.۳	تمرین‌ها	۱۵۰
۴	مدل‌های رگرسیونی خطی چندگانه	۱۵۷
۱.۴	مدل رگرسیون خطی چندگانه	۱۵۷
۲.۴	برآورد پارامترها به روش کم‌ترین توان دوم	۱۶۹
۳.۴	نمایش مدل رگرسیون خطی چندگانه با نماد ماتریسی	۱۷۲
۴.۴	استنباط درباره‌ی پارامترهای مدل رگرسیون خطی چندگانه	۱۸۸

۱۸۹	آزمون فرض و بازه‌ی اطمینان برای پارامترهای مدل رگرسیون خطی چندگانه	۱.۴.۴
۱۹۳	بازه‌ی اطمینان برای میانگین شرطی متغیر پاسخ	۲.۴.۴
۱۹۳	بازه‌ی پیش‌بینی	۳.۴.۴
۱۹۵	تحلیل واریانس در رگرسیون خطی چندگانه	۵.۴
۱۹۸	آزمون‌های جزئی ضریب‌های رگرسیونی	۶.۴
۲۰۲	آزمون فرض خطی کلی	۷.۴
۲۰۶	ضریب تعیین در مدل رگرسیون خطی چندگانه	۸.۴
۲۰۷	تقسیم‌هایی از مدل رگرسیون خطی چندگانه	۹.۴
۲۰۸	مدل رگرسیون چندجمله‌ای	۱۰.۴
۲۰۸	مدل رگرسیونی با اثرهای متقابل	۲.۴
۲۱۸	عیب‌یابی در مدل رگرسیون خطی چندگانه	۱۰.۴
۲۲۹	آشنایی با روش‌های تشخیصی	۱.۱۰.۴
۲۳۱	همبستگی چندگانه	۲.۱۰.۴
۲۳۶	تمرین‌ها	۱۱.۴
۲۴۷	انتخاب مدل	۵
۲۴۸	روش پیش‌رو	۱.۵
۲۴۹	ضریب همبستگی جزئی	۱.۱.۵
۲۵۱	ساختار روش پیش‌رو	۲.۱.۵
۲۵۷	روش‌های پس‌رو و گام به گام	۲.۵
۲۶۴	کنکاش بیشتر در مجموع توان دوم اضافی	۳.۵
۲۶۷	اثر همخطی چندگانه بر انتخاب مدل	۴.۵
۲۶۹	مشکلات تشخیص نادرست زیرمدل مناسب	۵.۵
۲۷۱	تمرین‌ها	۶.۵
۲۷۷	کتاب‌نامه	
۲۸۱	واژه‌نامه	
۲۸۵	نام‌نامه	
۲۸۷	نمایه	

پیش‌گفتار

از سال‌های دور، در پی آشنایی برای آماده‌سازی یک محتوای درخور برای درس رگرسیون بودم. نخستین فرصت در دانشگاه دکتی اسمهان فراهم شد، اما به بار ننشست. دومین فرصت در سال ۹۲ در دانشکده‌ی ریاضی، آمار و علوم کامپیوتر دانشگاه تهران فراهم شد.

نخستین قرار بر این بود که این کتاب توسط خانم دکتر افتخاری و این‌جانب نگاشته شود. در آغاز راه آقای دکتر امینی هم به جمع ما پیوستند و مسوولیت هماهنگی و بازنویسی و یک دست کردن مطالب در اختیار را که کاری بس طاقت‌فرسا است، بر عهده گرفتند. در میانه‌ی راه همکار دیگری از دانشگاه هرمزگان، آقای دکتر زمانی به جمع ما پیوستند و سبکی کار آقای دکتر امینی دوجندان شد.

امروزه، مدل‌های رگرسیونی در علوم اجتماعی، اقتصاد، پزشکی، سلامت، زیستی و ... کاربرد گسترده‌ای دارد. به‌کارگیری موفق این نوع مدل‌سازی نیازمند دانستن رگرسیون در دو زمینه‌ی نظری رگرسیون و چگونگی کاربرد قابل قبول آن در داده‌های واقعی است. کتاب پیش‌گفتار بر ارائه‌ی پایه‌ی نظری مدل‌سازی رگرسیونی در کنار رویکرد کاربردی این روش‌ها در حل مسائل واقعی است. در این راستا پس از بیان مباحث نظری در هر بخش، کدهای مربوط به نرم‌افزار برنامه‌نویسی R در مدل‌سازی داده‌های واقعی (یا داده‌های شبیه‌سازی) به همراه توضیحات ارائه می‌شود. هم‌چنین مجموعه‌ای از تمرین‌های نظری و عملی در انتهای هر فصل گنجانده شده است که خواننده را در ارزیابی میزان یادگیری و توجه بیش‌تر نسبت به مسائل پراهمیت پیرامون موضوع فصل یاری می‌نماید. گردآوری و آماده‌سازی مثال‌ها همراه با کدهای مربوط به نرم‌افزار برنامه‌نویسی R توسط آقای دکتر امینی فراهم شده و توسط خانم دکتر افتخاری مورد بازبینی قرار گرفته است.

کتاب پیش‌رو دست‌آورد تجربه‌ی تدریس همکاران مؤلف در درس رگرسیون ۱ دوره‌ی کارشناسی آمار است. تلاش شده است این کتاب یک منبع جامع در ارائه‌ی توأم مباحث نظری و کاربردی درس رگرسیون ۱ دوره کارشناسی رشته‌ی آمار و دیگر رشته‌های مرتبط باشد. مؤلفین بر این باور هستند که هر

فرد علاقه‌مند با داشتن آگاهی‌های پیش‌زمینه‌ای لازم در زمینه‌ی مبانی احتمال و آمار، یا احتمال ۱، ۲ و آمار ریاضی، می‌تواند به آسانی مطالب کتاب را دنبال کرده و حتی به عنوان یک کتاب خودآموز از آن بهره‌برد. این کتاب در ۵ فصل همراه با به‌کارگیری نرم‌افزار R آماده شده است و تدریس تمام مطالب آن به مدرسان گرامی توصیه می‌شود.

بر این باور هستیم که محتوای کتاب خالی از اشکال و ایراد نیست و هر نوع تذکر، اصلاح و یادآوری برای بهبود و بهتر شدن محتوای کتاب را از طرف دانشجویان و همکاران به دیده‌ی منت می‌پذیریم و ممنون و سپاس‌گزار این عزیزان خواهیم بود.

از آقای دکتر سید مرتضی امینی که زحمت حروف‌نگاری و یکسان‌سازی متن کتاب را به عهده داشتند، از خیم ذکر افتخاری که زحمت آماده‌سازی نهایی متن کتاب را به عهده داشتند و آقایان سید محمد خوش‌رو و یوسف حسینی که در آماده‌سازی کتاب ما را کمک کردند، آقای ناصر مجلسی راد که طراحی جلد کتاب را دنبال کردند و خانم مهدیه شهبازی که ویرایش، بازخوانی و صفحه‌آرایی کتاب را همراه با خانم دکتر افتخاری به عهده داشتند، تشکر و قدردانی می‌نماییم. بی‌شک اگر ثابت قدمی، پیگیری و تلاش‌های این عزیزان نبود، این اثر را اختیار خوانندگان قرار نمی‌گرفت.

از نظرات داور محترم کتاب که زحمت قبول داوری را به عهده گرفتند تشکر و قدردانی می‌نماییم. از مسئولین دانشگاه‌های تهران و هرمزگان به دلیل فراهم آوردن فضای مناسب برای تألیف و از انتشارات علمی پارسیان به دلیل حمایت و انتشار این کتاب تشکر ویژه داریم.

از طرف گروه مؤلفان

احمد پارسیان

بهار ۱۳۹۷