

تحلیل رگرسیون کاربردی

تألیف:

دکتر محمد تقی آساد

دکتر بهرام حیدری

اعضای هیئت علمی دانشگاه شیراز

آساد، محمد تقی -

تحلیل رگرسیون کاربردی / تألیف محمد تقی آساد - بهرام حیدری ؛ مشهد، جهاد دانشگاهی مشهد، ۱۳۹۰.

ص: مصور، جدول، نمودار. (انتشارات جهاد دانشگاهی مشهد ؛ ۴۳۲ : علوم پایه ؛ ۶۱).

ISBN 964-324-240-4

کتابنامه. ص: [۳۱۹] - ۳۲۰.

واژه نامه.

۱. برگشت (ریاضیات) الف. حیدری، بهرام، نویسنده همکار. ب. جهاد دانشگاهی مشهد.

ج. عنوان.

۵۳۶ / ۵۱۹

Q A ۲۷۸ / ۲ / ۲۲۵



انتشارات جهاد دانشگاهی مشهد

مشهد میدان آزادی، پردیس دانشگاه فردوسی، سازمان مرکزی جهاد دانشگاهی مشهد

ص: پ. ۱۳۷۶ - ۹۱۷۷۵ تلفن ۸۸۳۲۳۶۷ مرکز پخش ۸۸۴۲۲۳

E-mail: info@jdmpress.com

www.jdmpress.com

تحلیل رگرسیون کاربردی

تألیف: دکتر محمد تقی آساد - دکتر بهرام حیدری

حروفچینی و طراحی گرافیک: مشهد اسکندر چاپ و صحافی: دانشگاه فردوسی

چاپ اول پاییز ۱۳۹۰، ۱۱۰۰ نسخه، شماره نشر ۴۳۲

ISBN: 964-324-240-4

شابک ۹۶۴-۳۲۴-۲۴۰-۴

کلیه حقوق نشر برای ناشر محفوظ است.

قیمت: ۷۰۰۰۰ ریال

به نام خداوند جان و خرد

کتاب بزرگترین دستاورد فرهنگی بشر است. دانش بشری مدیون هزاران هزار کتابی است که در طول تاریخ با رنج و تلاش فراوان گرد آمده‌اند. کتاب تداوم معرفت علمی انسان است که سرانجام به تراکم دانش و بروز دگرگونی‌های تمدنی می‌انجامد.

جهاد دانشگاهی مشهد بر این باور است که نخستین گام در راه بهبود ساختارهای اقتصادی-اجتماعی و توسعه کشور، دستیابی به تازه‌های دانش و نشر یافته‌های پژوهشگران است. کتاب حاضر چهارصد و سی و دومین اثری است که با همین رویکرد منتشر می‌شود. رهنمودهای خوانندگان فرهیخته می‌تواند ما را در ارتقای سطح کیفی و کمی این آثار یاری نماید.

انتشارات جهاد دانشگاهی مشهد

فهرست

۱۱	پیش‌گفتار
۱۳	۱ مقدمه
۱۳	۱-۱ تعریف آمار
۱۵	۱-۲ مختصری از تاریخ آمار
۱۷	۱-۳ تحلیل رگرسیون
۲۰	منابع
۲۱	۲ جوامع معمول در پژوهش‌های زیستی
۲۱	۲-۱ مقدمه
۲۲	۲-۲ توزیع دو جمله‌ای
۲۴	۲-۳ توزیع پواسن
۲۵	۲-۴ توزیع نرمال
۲۸	دلیل استفاده از توزیع نرمال
۲۹	جداول توزیع نرمال
۳۳	انحراف معیار میانگین نمونه
۳۳	میانگین و انحراف معیار $\bar{X}$
۳۴	توزیع فراوانی میانگین نمونه‌ها
۳۵	حدود اطمینان برای μ هنگامی که σ معلوم باشد
۳۶	حدود اطمینان یک طرفه
۳۷	اندازه نمونه
۳۹	۲-۵ توزیع t استودنت
۴۰	حدود اطمینان μ در توزیع t

۴۲	۲-۶ تغییر نسبی - ضریب تغییر
۴۴	۲-۷ مقایسه میانگین دو نمونه
۴۵	آزمون معنی داری
۴۹	آزمون فرض برابری میانگین جمعیت با یک عدد خاص
۵۱	مقایسه میانگین دو نمونه مستقل با واریانس برابر (مشاهدات غیر جفتی)
۵۴	مقایسه میانگین های نمونه ای (مشاهدات جفتی)
۵۵	دستورالعمل آزمون t در SAS
۵۸	مقایسه میانگین نمونه های مستقل با واریانس های نابرابر
۶۰	۲-۸ آزمون فرض برابری بودن واریانس ها
۶۱	۲-۹ توان، اندازه نمونه و تشخیص اختلاف ها
۶۸	منابع

۶۹	۳ رگرسیون خطی
۶۹	۳-۱ مقدمه
۷۰	۳-۲ معادله یک خط راست
۷۰	۳-۳ رگرسیون Y روی X
۷۵	۳-۴ مدل رگرسیون خطی
۷۷	۳-۵ آزمون های آماری در رگرسیون
۷۷	قوانین امید ریاضی و واریانس
۸۲	۳-۶ تحلیل واریانس در رگرسیون خطی
۸۵	خط رگرسیونی که از مبدأ می گذرد
۸۶	دستورالعمل برازش رگرسیون ساده خطی در SAS
۸۹	۳-۷ آزمون عدم برازش مدل خطی
۹۱	دستورالعمل تعیین عدم برازش در SAS
۹۴	مسائل
۹۶	منابع

۹۷	۴ همبستگی خطی
۹۷	۴-۱ مقدمه
۹۷	۴-۲ همبستگی و ضریب آن
۹۹	محاسبه ضرایب همبستگی ساده در SAS

۱۰۱.....	ویژگی‌های ضریب همبستگی
۱۰۵.....	۴-۳ همبستگی و رگرسیون
۱۰۶.....	۴-۴ توزیع‌های نمونه‌ای، فاصله‌های اعتماد و آزمون فرض‌ها
۱۰۸.....	۴-۵ یکپارختی ضرایب همبستگی
۱۱۰.....	۴-۶ همبستگی درون‌گروهی
۱۱۳.....	مسائل
۱۱۴.....	منابع

۱۱۵.....	۵ عددنویسی با ماتریس
۱۱۵.....	۵-۱ مقدمه
۱۱۵.....	۵-۲ ماتریس‌ها
۱۱۶.....	۵-۳ عملیات ماتریس
۱۲۱.....	۵-۴ وارون، وابسته خطی و رتبه
۱۲۵.....	ریشه مشخصه و بردار مشخصه
۱۲۷.....	منابع

۱۲۸.....	۶ رگرسیون خطی با روش ماتریس
۱۲۸.....	۶-۱ مقدمه
۱۲۸.....	۶-۲ مدل و برآورد مجموع مربعات کمینه
۱۳۲.....	۶-۳ تحلیل واریانس
۱۳۳.....	۶-۴ انحراف معیارها، حدود اطمینان‌ها و آزمون فرض‌ها
۱۳۵.....	۶-۵ برآورد و پیش‌بینی
۱۳۶.....	مسائل
۱۳۸.....	منابع

۱۳۹.....	۷ رگرسیون خطی چندگانه و همبستگی جزئی
۱۳۹.....	۷-۱ مقدمه
۱۳۹.....	۷-۲ برآورد ضرایب
۱۴۱.....	۷-۳ مفهوم ضرایب رگرسیون
۱۴۲.....	مقایسه رگرسیون ساده و چندگانه
۱۴۳.....	حل مسائل رگرسیون چندگانه در SAS

۷-۴	حل مسائل رگرسیون خطی چندگانه با روش ماتریس	۱۴۷
۷-۵	ارزیابی مدل و آزمون‌های دیگر	۱۵۰
۷-۶	حالت متعامد	۱۵۴
۷-۷	همبستگی جزئی	۱۵۸
۷-۸	گزینش بهترین مدل رگرسیون	۱۵۹
	رگرسیون حذف پسر و در SAS	۱۶۲
	رگرسیون گام به گام	۱۶۵
	رگرسیون گام به گام در SAS	۱۶۷
	مسائل	۱۶۹
	منابع	۱۷۲

۸	رگرسیون غیر خطی	۱۷۳
۸-۱	مقدمه	۱۷۳
۸-۲	منحنی رشد نمایی	۱۷۵
۸-۳	چند جمله‌ای درجه دو	۱۷۷
	برآزش رگرسیون درجه ۲ در SAS	۱۸۲
	مسائل	۱۸۶
	منابع	۱۸۷

۹	بررسی شایستگی مدل رگرسیون	۱۸۸
۹-۱	مقدمه	۱۸۸
۹-۲	تجزیه باقی‌مانده‌ها (خطاها)	۱۹۰
	روش‌های نشان‌دادن باقی‌مانده‌ها	۱۹۰
۹-۳	روش رسم شکل	۱۹۷
۹-۴	رسم شکل جهت بررسی نرمال بودن	۱۹۸
	رسم باقی‌مانده‌ها نسبت به مقادیر برآزش شده $\hat{y}$	۲۰۰
	رسم داده‌ها در برابر مانده‌ها در SAS	۲۰۳
	آزمون نرمال برای داده‌ها در SAS	۲۰۷
۹-۵	تبدیل متغیرها	۲۱۵
	تبدیل برای تثبیت واریانس	۲۱۵
	تبدیل برای خطی کردن مدل	۲۱۹

۲۲۵	تبدیل لگاریتمی متغیر مستقل (X) در SAS
۲۲۷	تبدیل لگاریتمی متغیر وابسته (Y) در SAS
۲۲۸	۹-۶ متغیرهای نشانگر
۲۲۹	متغیر ظاهری با بیش از دو سطح
۲۳۷	۹-۷ مجموع مربعات اضافی
۲۴۰	تجزیه SSR به مجموع مربعات اضافی
۲۴۲	۹-۸ چند هم خطی و اثرهای آن
۲۴۳	متغیرهای مستقل همبستگی ندارند
۲۴۴	متغیرهای مستقل کاملاً وابسته اند
۲۴۶	آثار چند هم خطی
۲۴۷	تأثیر بر ضرایب رگرسیونی
۲۴۷	تأثیر بر مجموع مربعات اضافی
۲۴۸	تأثیر بر $S(b_k)$
۲۴۸	تأثیر بر مقادیر پیش بینی شده $(\hat{Y}_k)$ و میان تغییرات آن $(S\hat{Y}_k)$
۲۴۹	تأثیر بر آزمون های همزمان β_k
۲۵۰	۹-۹ تشخیص چند هم خطی
۲۵۴	نحوه مقابله با چند هم خطی
۲۵۸	۹-۱۰ مجموع مربعات موزون
۲۶۳	مسایل
۲۶۷	منابع
۲۶۸	۱۰ تجزیه ضرایب مسیر
۲۶۸	۱۰-۱ مقدمه
۲۶۹	۱۰-۲ ضرایب رگرسیون ساده استاندارد
۲۶۹	۱۰-۳ مقیاس بندی واحد نرمال
۲۷۰	۱۰-۴ مقیاس بندی واحد طول
۲۷۰	۱۰-۵ ضرایب رگرسیون جزء استاندارد
۲۷۲	۱۰-۶ علت و معلولی و ضرایب مسیر
۲۷۳	۱۰-۷ ضرایب مسیر و تبیین
۲۷۸	ضرایب مسیر برای دو متغیر وابسته (معلول)
۲۸۴	مسایل
۲۹۰	منابع

۲۹۱.....	۱۱ تجزیه کوواریانس
۲۹۱.....	۱۱-۱ مقدمه
۲۹۲.....	۱۱-۲ کوواریانس و رابطه بین متغیرها
۲۹۲.....	۱۱-۳ کاربردهای تجزیه کوواریانس
۲۹۳.....	۱۱-۴ طرح های دارای یک کوواریت
۳۰۲.....	تجزیه کوواریانس برای مثال ۱ به کمک نرم افزار SAS
۳۰۴.....	۱۱-۵ مدل های دیگر کوواریانس
۳۰۶.....	۱۱-۶ تجزیه کوواریانس در آزمایش های کاملاً تصادفی
۳۰۹.....	۱۱-۷ تجزیه کوواریانس در طرح های با طبقه بندی یک طرفه
۳۱۲.....	۱۱-۸ تجزیه کوواریانس در طرح های با طبقه بندی دوطرفه
۳۱۳.....	۱۱-۹ تجزیه کوواریانس چندگانه
۳۱۶.....	مسائل
۳۱۸.....	منابع
۳۱۹.....	کل منابع
۳۲۱.....	پیوست ها
۳۲۳.....	نمایه

پیشگفتار

تجزیه و تحلیل رگرسیون یکی از مهمترین روش‌های تحلیل داده‌های چند متغیره است. مزیت و استفاده گسترده از تحلیل‌های رگرسیون ناشی از قابلیت یک معادله در بیان روابط بین یک متغیر پاسخ و مجموعه‌ای از متغیرهای تأثیرگذار بر آن است. آگاهی از تئوری مدل‌های رگرسیون و همچنین توجه به مشکلات کاربردی که حین به کارگیری آن برای تحلیل داده‌های واقعی به وجود می‌آید ضروری بوده و از اهمیت زیادی برخوردار است. اگرچه اهداف، روش‌ها و استنتاج‌های آماری مورد نیاز پژوهش‌گران اغلب ثابت و بدون تغییر است ولی معرفی و طراحی نرم‌افزارهای آماری و استفاده از کامپیوتر این امکان را به پژوهش‌گران داده است تا پیچیده‌ترین روش‌های آماری را به راحتی برای داده‌های خود به کار ببرند. در برخی موارد مشاهده می‌شود که برخی از دانشجویان بدون توجه به طبیعت داده‌های خود مدل‌های رگرسیونی را به کار می‌برند که ممکن است به نتایج نادرست منتج شود. در این کتاب چگونگی تجزیه و تحلیل صحیح داده‌ها و همچنین تفسیر نتایج به دست آمده از نرم‌افزار SAS به همراه مثال‌های متعدد در هر بخش شرح داده می‌شود. در پایان هر بخش مسایل متعدد و اغلب همراه با جواب آمده است تا محکی برای میزان فراگیری دانشجویان باشد. آشنایی با مباحث آمار مقدماتی نظیر انواع توزیع‌های فراوانی و آماره‌های توصیفی مانند میانگین و واریانس اساس و پیش زمینه درک و یادگیری بحث‌های پیچیده‌تر در آمار پیشرفته را فراهم می‌سازد. تجربه نگارندگان نشان می‌دهد که بعضی از دانشجویان در این زمینه آگاهی کافی ندارند. فصل ۲ کتاب به معرفی انواع جوامع معمول در پژوهش‌های رشته‌های مختلف علوم، مقدمات آمار و مثال‌های مرتبط اختصاص یافته است. در فصل‌های ۳ و ۴ اصول و مبانی رگرسیون ساده خطی و همبستگی بین متغیرها به بحث گذاشته شده است. آشنایی با جبر ماتریس‌ها و چگونگی حل مسایل رگرسیون ساده خطی با روش ماتریس در فصل‌های ۵ و ۶ آورده شده است. در فصل ۷، رگرسیون چندگانه و روش‌های گزینش بهترین مدل رگرسیون بر مبنای چند متغیر مستقل به همراه چند مثال حل شده در نرم افزار SAS به بحث گذاشته شده است. در فصل ۸ انواع مدل‌های غیرخطی و چگونگی برازش آنها آمده است. خوانندگان می‌توانند برای آشنایی بیشتر با روش‌های بررسی شایستگی مدل‌های رگرسیون،

چگونگی آزمون داده‌های نرمال و تبدیل داده‌ها به کمک نرم‌افزار، آثار متغیرهای همبسته بر مدل‌های رگرسیون و تشخیص هم‌خطی به فصل ۹ مراجعه نمایند. در فصل‌های ۱۰ و ۱۱ کاربرد تجزیه ضرایب مسیر و تجزیه کوواریانس که همواره کمتر به آن توجه شده است با ذکر مثال توضیح داده شده است. به نظر می‌رسد که در بین کتاب‌های فارسی کتابی با این محتوا و نگرش وجود ندارد، لذا نگارندگان امیدوارند این کتاب بتواند نیاز دانشجویان کارشناسی و تحصیلات تکمیلی و پژوهش‌گران جوان در رشته‌های مختلف علوم از جمله زیست‌شناسی، کشاورزی، پزشکی، دامپزشکی، اقتصاد، جامعه‌شناسی و غیره را برآورده سازد. در مواردی که خواننده نیاز به کسب اطلاعات بیشتری دارد می‌تواند به منابع ارائه شده در انتهای هر فصل مراجعه نماید.

نگارندگان

assad@shirazu.ac.ir

bheidari@shirazu.ac.ir

www.ketab.ir