

برنام‌خواندگان

روش‌های پیشرفته آماری و مدل‌های خطی در علوم دامی



با استفاده از نرم‌افزار R

دکتر مهدی جباری نوقابی

عضو هیئت علمی دانشگاه فردوسی مشهد

دکتر خدیجه نصیری

عضو هیئت علمی دانشگاه مازندران

دکتر زهرا رودباری

عضو هیئت علمی دانشگاه جیرفت

دکتر مرتضی بیطرف ثانی

عضو هیئت علمی مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان یزد

دکتر مهدی خجسته‌کی

عضو هیئت علمی مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان قم

سرشناسه: عنوان و نام پدیدآور: جباری نوقایی، مهدی، ۱۳۵۷ - روش‌های پیشرفته آماری و مدل‌های خطی در علوم دامی (با استفاده از نرم‌افزار R) / تألیف مهدی جباری نوقایی و همکاران خدیجه نصیری ... او دیگران؛ ویراستار علمی بهرام صادق‌پورگیله.

مشخصات نشر: مشهد: دانشگاه فردوسی مشهد، انتشارات، ۱۳۹۵.

مشخصات ظاهری: ۴۴۴ ص.: مصور، جدول، نمودار.

فروست: انتشارات دانشگاه فردوسی مشهد؛ ۶۵۹.

شابک: ISBN: 978-964-386-334-0

وضعیت فهرست‌نویسی: فایا.

یادداشت: همکاران خدیجه نصیری، زهرا رودباری، مرتضی بیطرف ثانی، مهدی خجسته‌کی.

یادداشت: عنوان اصلی: M. Jabbari Nooghahi ... [et. al]. Advanced statistical methods and statistical models in animal sciences (with application in R).

یادداشت: واژه‌نامه. کتابنامه: ص. [۴۳۵] - ۴۳۸.

یادداشت: چاپ دوم: ۱۴۰۴ (فیبا).

موضوع: دامپروری -- روش‌های آماری Animal culture -- Static methods
دامپزشکی -- روش‌های آماری Veterinary medicine -- Statistical methods
آر (زبان برنامه‌نویسی کامپیوتر) R (Computer program language)

شناسه افزوده: نصیری، خدیجه، ۱۳۶۴ -

شناسه افزوده: دانشگاه فردوسی مشهد، انتشارات.

رده‌بندی کنگره: ۱۳۹۵ ج۲ر۹ / SF۱۴۰

رده‌بندی دیویی: ۶۳۶/۰۰۷۲۷

شماره کتابشناسی ملی: ۴۲۲۴۱۰۴

روش‌های پیشرفته آماری و مدل‌های خطی در علوم دامی

(با استفاده از نرم‌افزار R)

پدیدآورندگان: دکتر مهدی جباری نوقایی؛ دکتر خدیجه نصیری؛ دکتر زهرا رودباری
دکتر مرتضی بیطرف ثانی؛ دکتر مهدی خجسته‌کی
ویراستار علمی: دکتر بهرام صادق‌پورگیله
مشخصات: وزیری، ۱۵۰ نسخه، چاپ دوم، زمستان ۱۴۰۴ (اول، ۱۳۹۵)
چاپ و صحافی: همیار
بها: ۷/۱۰۰/۰۰۰ ریال
حق چاپ برای انتشارات دانشگاه فردوسی مشهد محفوظ است.



انتشارات
۶۵۹

مراکز بخش:

فروشگاه و نمایشگاه کتاب پردیس: مشهد، میدان آزادی، دانشگاه فردوسی مشهد، جنب سلف یاس
تلفن: ۳۸۸۰۲۶۶۶ - ۳۸۸۳۳۷۲۷ (۰۵۱)

مؤسسه کتابیران: تهران، میدان انقلاب، خیابان کارگر جنوبی، بین روانمهر و وحید نظری، بن‌بست
گشتاسب، پلاک ۸ تلفن: ۶۶۴۸۴۷۱۵ (۰۲۱)

مؤسسه دانشوران: تهران، خیابان انقلاب، خیابان منیری جاوید (اردیبهشت) نبش خیابان نظری، شماره ۱۴۲
تلفکس: ۶۶۴۰۰۲۲۰ - ۶۶۴۰۰۱۴۴ (۰۲۱)

<http://press.um.ac.ir>

Email: press@um.ac.ir

فهرست

پیشگفتار.....	۱۱
فصل اول: مقدمه‌ای بر نرم افزار آماری R.....	۱۳
۱-۱- نرم افزار آماری R.....	۱۳
۱-۲- نحوه نصب نرم افزار R.....	۱۴
۱-۳- مزایای نرم افزار R.....	۱۴
۱-۴- محدودیتهای نرم افزار R.....	۱۵
۱-۵- آشنایی با متوهای نرم افزار R.....	۱۵
۱-۵-۱- منوی File.....	۱۶
۱-۵-۲- منوی Edit.....	۱۶
۱-۵-۳- منوی Misc.....	۱۶
۱-۵-۴- منوی Packages.....	۱۶
۱-۵-۴-۱- نصب و فراخوانی بسته‌ها.....	۱۷
۱-۵-۵- منوی Help.....	۱۸
۱-۵-۶- منوی Windows.....	۱۹
۱-۶- شروع تحلیل آماری برای یک پروژه با نرم افزار R.....	۱۹
۱-۷- نمایش دستورات در نرم‌افزار R.....	۱۹
۱-۸- مزایای پنجره R script یا ویرایشگر متنی.....	۲۰
۱-۹- ایجاد، فهرست و حذف کردن اشیاء در حافظه.....	۲۰
۱-۹-۱- عملگرهای حسابی.....	۲۱
۱-۹-۲- عملگرهای نسبی و منطقی.....	۲۱
۱-۱۰- داده‌ها در نرم افزار R.....	۲۱
۱-۱۱- خواندن داده‌ها از فایل.....	۲۳
۱-۱۱-۱- دستور خواندن read.table.....	۲۳
۱-۱۱-۱-۱- دستور خواندن فایل‌های نرم‌افزار Excel.....	۲۴
۱-۱۱-۲- دستور خواندن scan.....	۲۵
۱-۱۲- تولید داده‌ها.....	۲۵
۱-۱۲-۱- ایجاد دنباله منظم از داده‌ها.....	۲۵

۴ روش‌های پیشرفته آماری و مدل‌های خطی در علوم دامی

- ۲۶-۱۲-۲- ایجاد دنباله تصادفی از داده‌ها.....
- ۲۷-۱- برنامه نویسی با نرم افزار R.....
- ۲۹- روشهای پیشرفته آماری.....
- فصل دوم: مفاهیم آماری..... ۳۱**
- ۳۱-۱-۲- آمار.....
- ۳۱-۲-۲- علم آمار.....
- ۳۲-۳-۲- روشهای آماری.....
- ۳۲-۱-۳-۲- آمار توصیفی.....
- ۳۲-۲-۳-۲- آمار استنباطی.....
- ۳۲-۳-۳-۲- روشهای آماری ناپارامتری (غیر پارامتریک).....
- ۳۳-۴-۲- جامعه آماری.....
- ۳۳-۵-۲- نمونه آماری.....
- ۳۳-۶-۲- متغیر آماری.....
- ۳۴-۱-۶-۲- متغیر کمی.....
- ۳۴-۲-۶-۲- متغیر کیفی.....
- ۳۵-۷-۲- مقیاس اندازه گیری.....
- ۳۵-۱-۷-۲- انواع مقیاس اندازه گیری.....
- ۳۷-۸-۲- متغیر مستقل و وابسته (نقش یک متغیر).....
- ۳۸-۹-۲- شاخص‌های تمایل مرکزی.....
- ۳۸-۱-۹-۲- مُد.....
- ۳۹-۲-۹-۲- میانه.....
- ۴۰-۳-۹-۲- میانگین حسابی.....
- ۴۰-۴-۹-۲- میانگین هندسی.....
- ۴۰-۵-۹-۲- میانگین همساز یا هارمونیک.....
- ۴۱-۱۰-۲- شاخص‌های پراکندگی.....
- ۴۲-۱-۱۰-۲- دامنه تغییرات.....
- ۴۲-۲-۱۰-۲- دامنه میان چارکی.....
- ۴۲-۳-۱۰-۲- میانگین قدر مطلق انحرافات از میانگین حسابی.....
- ۴۳-۴-۱۰-۲- واریانس.....
- ۴۳-۵-۱۰-۲- انحراف معیار.....
- ۴۳-۱۱-۲- دستور برخی توابع پایه آماری در نرم افزار R.....
- ۴۴-۱۲-۲- انواع مدل‌های آماری.....
- ۴۶-۱۳-۲- برخی از تفاوت‌های تحلیل واریانس و کوواریانس.....

فهرست ۵

۴۷	۱۴-۲- همبستگی و انواع آن
۴۷	۱-۱۴-۲- همبستگی
۴۹	۲-۱۴-۲- انواع ضرایب همبستگی
۵۰	۱-۲-۱۴-۲- حالتی که دو متغیر اسمی و یا یکی اسمی و دیگری رتبه‌ای باشد
۵۲	۲-۲-۱۴-۲- ضریب همبستگی در حالتی که دو متغیر دارای مقیاس رتبه‌ای باشند
۵۳	۳-۱۴-۲- متغیر اول با مقیاس اسمی یا رتبه‌ای و متغیر دیگر با مقیاس فاصله‌ای یا نسبی
۵۴	۴-۱۴-۲- ضریب همبستگی در حالتی که دو متغیر مقیاس فاصله‌ای یا نسبی داشته باشند
۵۹	۱۵-۲- تمرین
۶۱	فصل سوم: جبر ماتریسها
۶۱	۱-۳- مقدمه
۶۳	۲-۳- انواع ماتریسها
۶۵	۳-۳- عملیات روی ماتریسها
۷۸	۴-۳- تمرین
۷۹	فصل چهارم: برآورد فاصله‌های و آزمون فرضیه آماری
۷۹	۱-۴- برآورد نقطه‌ای برای میانگین جمعیت
۸۰	۲-۴- برآورد فاصله‌ای (فاصله اطمینان)
۸۰	۱-۲-۴- برآورد حدود اطمینان میانگین جمعیت با واریانس معلوم
۸۱	۲-۲-۴- برآورد حدود اطمینان میانگین جمعیت با واریانس نامعلوم
۸۲	۳-۴- برآورد حجم نمونه در برآورد فاصله‌ای برای میانگین جمعیت
۸۳	۴-۴- برآورد نقطه‌ای نسبت جمعیت
۸۳	۵-۴- برآورد فاصله‌ای نسبت جمعیت
۸۴	۶-۴- تعیین اندازه نمونه در برآورد فاصله‌ای برای نسبت جمعیت
۸۵	۷-۴- آزمون فرضیه
۸۶	۱-۷-۴- انواع فرضیه‌ها
۸۷	۸-۴- انواع خطاها
۸۷	۹-۴- آماره آزمون
۸۸	۱۰-۴- ناحیه بحرانی
۸۸	۱۱-۴- سطح معنی داری آزمون
۸۸	۱۲-۴- آزمون‌های یک طرفه و دو طرفه
۸۹	۱۳-۴- آزمون فرضیه یک طرفه چپ میانگین جمعیت با واریانس معلوم
۹۰	۱۴-۴- آزمون یک طرفه دامنه بالای میانگین جمعیت با واریانس معلوم
۹۱	۱۵-۴- آزمون دو طرفه میانگین جمعیت با واریانس معلوم

۶ روش‌های پیشرفته آماری و مدل‌های خطی در علوم دامی

- ۹۲-۴-۱۶-آزمون فرضیه در جوامع با واریانس نامعلوم.....
- ۹۴-۴-۱۷-آزمون یک طرفه دامنه پایین نسبت جمعیت.....
- ۹۶-۴-۱۸-آزمون یک طرفه دامنه بالایی نسبت جمعیت.....
- ۹۷-۴-۱۹-آزمون فرضیه دو طرفه نسبت جمعیت.....
- ۹۸-۴-۲۰-خطای نوع دوم.....
- ۹۸-۴-۲۰-۱-برآورد خطای نوع دوم در آزمون یک طرفه دامنه پایین میانگین جمعیت با واریانس معلوم.....
- ۱۰۰-۴-۲۰-۲-برآورد خطای نوع دوم در آزمون یک طرفه دامنه بالای میانگین جمعیت با واریانس معلوم.....
- ۱۰۱-۴-۲۰-۳-برآورد خطای نوع دوم در آزمون دو طرفه میانگین جمعیت با واریانس معلوم.....
- ۱۰۲-۴-۲۰-۴-برآورد خطای نوع دوم در آزمون یک طرفه دامنه پایین میانگین جمعیت با واریانس نامعلوم.....
- ۱۰۳-۴-۲۰-۵-برآورد خطای نوع دوم در آزمون یک طرفه دامنه بالای میانگین جمعیت با واریانس نامعلوم.....
- ۱۰۴-۴-۲۰-۶-برآورد خطای نوع دوم در آزمون دو طرفه میانگین جمعیت با واریانس نامعلوم.....
- ۱۰۵-۴-۲۱-روش برآورد پارامترها.....
- ۱۰۶-۴-۲۲-تابع درستمایی.....
- ۱۱۰-۴-۲۳-تمرین.....

۱۱۱ فصل پنجم: رگرسیون (خطی، غیر خطی و چندگانه).....

- ۱۱۱-۵-۱-رگرسیون.....
- ۱۱۲-۵-۲-رگرسیون خطی.....
- ۱۱۴-۵-۳-مدل رگرسیون خطی ساده.....
- ۱۱۵-۵-۳-۱-پذیره‌های زیر بنایی مدل.....
- ۱۱۶-۵-۳-۲-ویژگی‌های مدل.....
- ۱۱۷-۵-۴-برآورد پارامترهای رگرسیون-برآورد کمترین توانهای دوم.....
- ۱۲۳-۵-۵-رگرسیون خطی چندگانه.....
- ۱۲۹-۵-۶-رگرسیون چند جمله‌ای.....
- ۱۳۲-۵-۷-رگرسیون غیر خطی.....
- ۱۳۳-۵-۸-رگرسیون لگاریتمی.....
- ۱۳۳-۵-۸-۱-مقدمه.....
- ۱۳۷-۵-۹-تمرین.....

۱۴۱ فصل ششم: رگرسیون با متغیر پاسخ گسسته.....

- ۱۴۱-۶-۱-رگرسیون.....
- ۱۴۲-۶-۲-تعریف رگرسیون لوژستیک.....
- ۱۴۴-۶-۳-چند مثال از کاربردهای رگرسیون لوژستیک.....
- ۱۴۵-۶-۴-مزایای استفاده از رگرسیون لوژستیک.....

۷ فهرست

۱۴۵	۵-۶- معایب رگرسیون لوژستیک.....
۱۴۶	۶-۶- مفهوم تابع و مدل لوژستیک.....
۱۴۷	۷-۶- برآورد ماکسیمم درستمایی.....
۱۴۸	۸-۶- فاصله اطمینان برای نسبت بخت.....
۱۴۸	۹-۶- آزمون فرضیه در مدل‌های خطی تعمیم یافته.....
۱۴۹	۱۰-۶- آزمون هوسمر-لمشو.....
۱۴۹	۱۱-۶- آزمون معنیداری ضرایب رگرسیون.....
۱۵۰	۱۲-۶- رگرسیون لوژستیک چند سطحی.....
۱۵۱	۱۳-۶- روش انجام رگرسیون لوژستیک.....
۱۵۵	۱۴-۶- روش انجام رگرسیون لوژستیک چند سطحی در نرم افزار R.....
۱۶۲	۱۵-۶- رگرسیون پروبیت.....
۱۶۲	۱۶-۶- تابع پیوند.....
۱۶۳	۱۷-۶- مدل پروبیت.....
۱۶۴	۱۸-۶- مقایسه مدل لجیت و پروبیت.....
۱۶۷	۱۹-۶- رگرسیون لوژستیک ترتیبی.....
۱۶۸	۲۰-۶- مدل رگرسیون ترتیبی.....
۱۷۱	۲۱-۶- رگرسیون متعامد.....
۱۷۱	۲۲-۶- مفهوم باقیمانده مدل (خطای مدل).....
۱۷۲	۲۳-۶- تعریف دیگری از باقیمانده (خطای مدل).....
۱۷۳	۲۴-۶- روش کمترین توانهای دوم معمولی.....
۱۷۴	۲۵-۶- فرضهای کلاسیک (پذیرهای زیربنایی مدل).....
۱۷۴	۲۶-۶- نقض فرضهای کلاسیک.....
۱۷۷	۲۷-۶- نحوه محاسبه خطاها در رگرسیون متعامد.....
۱۷۹	۲۸-۶- برآورد ضرایب مدل رگرسیون متعامد با استفاده از روش کمترین توانهای دوم دو مرحله‌ای.....
۱۷۹	۲۹-۶- تعریف متغیر ابزاری.....
۱۸۵	۳۰-۶- برخی مشکلات استفاده از متغیر ابزاری.....
۱۸۶	۳۱-۶- تعیین میزان قدرت متغیرهای ابزاری.....
۱۸۷	۳۲-۶- تمرین.....
۱۹۳	فصل هفتم: آزمونهای آماری در مدل رگرسیون خطی.....
۱۹۳	۱-۷- بررسی پذیره‌های زیر بنایی مدل.....
۱۹۴	۲-۷- مانده‌ها.....
۱۹۴	۳-۷- بررسی برازش توزیع نرمال به مانده‌های مدل.....
۲۰۱	۴-۷- عدم خود همبستگی مرتبه یک مانده‌ها.....

۸ روش‌های پیشرفته آماری و مدل‌های خطی در علوم دامی

- ۲۰۲-۵-۷- روشهای بررسی عدم خود همبستگی مرتبه یک مانده‌ها.....
- ۲۰۲-۱-۵-۷- آزمون دوربین-واتسون.....
- ۲۰۳-۱-۵-۷- طریقه محاسبه آماره آزمون دوربین-واتسون با استفاده از نرم افزار R.....
- ۲۰۶-۲-۱-۵-۷- محدودیت‌های آزمون دوربین-واتسون.....
- ۲۰۶-۲-۵-۷- آزمون گشتها.....
- ۲۰۷-۳-۵-۷- نمودار مانده‌ها در مقابل ترتیب زمان (شماره مشاهدات).....
- ۲۰۹-۶-۷- واریانس مانده‌ها ثابت باشد.....
- ۲۰۹-۱-۶-۷- نمودار مانده‌ها در مقابل مقادیر برازش شده.....
- ۲۱۱-۲-۶-۷- نمودار مانده‌ها در مقابل مقادیر متغیرهای پیشگو.....
- ۲۱۱-۳-۶-۷- بررسی ثبات واریانس مانده‌ها با استفاده از نرم افزار R.....
- ۲۱۳-۷-۷- تبدیلات تثبیت کننده واریانس مانده‌ها در مدل رگرسیون خطی.....
- ۲۱۴-۱-۷-۷- تبدیل عمومی باکس و کاکس روی λ
- ۲۱۸-۸-۷- همخطی چندگانه.....
- ۲۲۲-۹-۷- مشاهدات کرانهای.....
- ۲۲۵-۱۰-۷- شناسایی مشاهدات کرانهای با نرم افزار R.....
- ۲۲۹-۱۱-۷- دستورهای رسم نمودارهای تشخیصی مدل.....
- ۲۳۴-۱۲-۷- پیشبینی برای آینده با دستور predict در نرم افزار R.....
- ۲۳۴-۱۳-۷- شاخص PRESS.....
- ۲۳۶-۱۴-۷- معیارهای انتخاب بهترین مدل.....
- ۲۳۶-۱-۱۴-۶- ضریب تعیین مدل.....
- ۲۳۷-۲-۱۴-۷- میانگین توان دوم مانده‌ها.....
- ۲۳۷-۳-۱۴-۷- معیار CP.....
- ۲۳۸-۴-۱۴-۷- معیار آکائیک.....
- ۲۳۹-۱۵-۷- فنونی محاسباتی برای انتخاب متغیر.....
- ۲۳۹-۱-۱۵-۷- روش همه رگرسیونهای ممکن.....
- ۲۳۹-۲-۱۵-۷- روشهای رگرسیونی قدم به قدم.....
- ۲۵۳-۱۶-۷- کمبود برازش و رگرسیون لگاریتمی.....
- ۲۵۳-۱-۱۶-۷- مقدمه.....
- ۲۵۵-۱۷-۷- آزمون کمبود برازش با واریانس معلوم.....
- ۲۵۵-۱۸-۷- آزمون کمبود برازش با واریانس نامعلوم.....
- ۲۶۱-۱۹-۷- تحلیل حساسیت مدل.....
- ۲۶۱-۱-۱۹-۷- تحلیل حساسیت.....
- ۲۷۰-۲۰-۷- تمرین.....

فصل هشتم: بهترین پیشبینی نااریب خطی ارزشهای اصلاحی و مدل دام.....	۲۷۵
۱-۸- مقدمه.....	۲۷۵
۲-۸- مدل های خطی.....	۲۷۶
۳-۸- روشهای تشخیص عوامل ثابت و تصادفی.....	۲۷۷
۴-۸- معادلات مدل آمیخته (مختلط).....	۲۷۷
۵-۸- استفاده از معادلات مدل آمیخته برای برآورد اثرات ثابت و پیش بینی ارزشهای اصلاحی با استفاده از مدل دام.....	۲۸۰
۶-۸- مدل حیوانی با گروه بندی.....	۲۸۸
۱-۶-۸- گروه بندی حیوانات.....	۲۸۸
۷-۸- تمرین.....	۲۹۴
فصل نهم: برآورد اثرات ژنتیکی مادری در مدل دام.....	۲۹۷
۱-۹- مقدمه.....	۲۹۷
۲-۹- اثرات ژنتیکی مادری در انواع حیوانات.....	۲۹۸
۳-۹- صفات مهم تحت تأثیر اثرات ژنتیکی مادری.....	۲۹۹
۴-۹- نحوه برآورد اثرات ژنتیکی مادری.....	۲۹۹
۵-۹- مفهوم مدل دام.....	۳۰۰
۶-۹- مدل دام با اثرات ژنتیکی مادری.....	۳۰۰
۷-۹- ساختار دادهها برای برآورد اثرات ژنتیکی مادری.....	۳۰۱
۸-۹- تمرین.....	۳۱۴
فصل دهم: برآورد اثر ژنتیکی و محیطی مادری با استفاده از مدل دام.....	۳۱۷
۱-۱۰- مقدمه.....	۳۱۷
۲-۱۰- مدل دام با اثر ژنتیکی و محیطی مادری.....	۳۱۹
۱-۲-۱۰- اثر ژنتیکی افزایشی مادری.....	۳۲۰
۲-۲-۱۰- از طریق توارث سیتوپلاسمی.....	۳۲۰
۳-۲-۱۰- اثر محیطی مادری.....	۳۲۰
۳-۱۰- مدل دام با اثر ژنتیکی و محیطی دائمی و مشترک مادری.....	۳۲۲
حل مثال ۱-۱۰ با استفاده از نرم افزار R.....	۳۲۵
۴-۱۰- تمرین.....	۳۳۲
فصل یازدهم: برآورد اثرات ژنتیکی مدل دام چند صفتی.....	۳۳۵
۱-۱۱- مقدمه.....	۳۳۵
۲-۱۱- مدل اثرات ژنتیکی چند صفتی.....	۳۳۷
۳-۱۱- تشکیل ماتریسهای طرح، ماتریس خویشاوندی و معکوس ماتریس خویشاوندی.....	۳۳۸

۳۴۰ محاسبه ماتریسهای X_1 و X_2
۳۴۱ محاسبه ماتریسهای Z_1 و Z_2
۳۴۲ تعیین ضرایب ماتریسهای G و R
۳۴۷ ۴-۱۱- تمرین
۳۴۹	فصل دوازدهم: برآورد اثرات ژنتیکی غیر افزایشی
۳۴۹ ۱-۱۲- اثرات غیر افزایشی
۳۵۰ ۲-۱۲- ماتریس روابط خویشاوندی غالبیت
۳۵۰ ۳-۱۲- مدل دام با اثر غالبیت
۳۵۱ ۴-۱۲- مدل آمیخته برای برآورد اثرات ثابت، افزایشی و غالبیت
۳۶۶ ۵-۱۲- اثرات اپیستازی
۳۶۷ ۶-۱۲- مدل آمیخته برای برآورد اثرات ثابت، افزایشی، غالبیت و اثرمتقابل اپیستازی افزایشی \times غالبیت
۳۸۸ ۷-۱۲- تمرین
۳۹۱	فصل سیزدهم: رگرسیون تصادفی
۳۹۱ ۱-۱۳- رکوردهای تکرار شده در طی زمان
۳۹۲ ۱-۱-۱۳- مدل همگن
۳۹۲ ۲-۱-۱۳- مدل ناهمگن
۳۹۳ ۲-۱۳- رگرسیون با ضرایب تصادفی
۳۹۳ ۳-۱۳- کاربرد مدل رگرسیون تصادفی در تجزیه و تحلیل داده‌های طولی در اصلاح نژاد دام و طیور
۳۹۴ ۴-۱۳- مدل رکوردهای روز آزمون
۳۹۴ ۵-۱۳- ساختار مدل رگرسیون تصادفی
۳۹۵ ۶-۱۳- توابع متعامد
۳۹۶ ۷-۱۳- چند جمله‌ای‌های لژاندر
۳۹۷ ۸-۱۳- تابع کوواریانس
۴۰۰ ۹-۱۳- شکل ماتریسی مدل رگرسیون تصادفی
۴۰۰ ۱۰-۱۳- ماتریس کوواریانس اثرات تصادفی
۴۲۰ ۱۱-۱۳- تمرین
۴۴۱ پیوست‌ها
۴۳۵ منابع
۴۳۹ واژه نامه انگلیسی به فارسی

پیشگفتار

در روزگار امروز همه پژوهشگران برای انجام تحقیقات نیاز به استفاده از روش‌های آماری دارند. در سال‌های اخیر به همت تلاش‌های متخصصین آماری و نیز علاقه‌مندان علم آمار استفاده از این روش‌ها بسیار متداول شده و به صورت عمومی تقریباً تمامی دانش‌پژوهان از روش‌های آماری اطلاع دارند. در این بین، نکته مهم آن است که این پژوهش‌گران تنها اطلاعات لایه رویی روش‌های آماری را می‌دانند و به زوایای پنهان آن آشنا نیستند. در این کتاب سعی شده است که با توجه به خصوصیات روش‌های آماری پیشرفته این نکات مهم به صورت کاربردی مطرح و با استفاده از مثال‌های واقعی و با به کار بردن نرم افزار آماری R نتایج تجزیه و تحلیل‌های آماری تشریح شود. نرم افزار آماری R یکی از نرم افزارهای منبع باز و قدرتمند آمار است که با استفاده از آن می‌توان تقریباً تمامی روش‌های آماری را انجام داد. برخی از این روش‌ها قبلاً توسط سایر پژوهشگران آمار مورد بررسی قرار گرفته و برنامه‌ها و دستورات مربوط به آن‌ها تحت عنوان بسته‌های مختلف تهیه و در مخازن این نرم افزار به صورت رایگان در اختیار قرار گرفته شده است. اما برخی دیگر از روش‌ها هنوز جای کار بسیاری را می‌طلبند. از طرفی، ممکن است یک پژوهشگر روش آماری مورد استفاده خود را بداند اما نتواند از این نرم افزار استفاده کند، لذا در این کتاب سعی شده است که با این رویکرد، علاقه‌مندان راهنمایی شوند.

بنابراین، این کتاب می‌تواند یک مرجع مناسب برای دانشجویان کارشناسی و کارشناسی ارشد آمار، دانشجویان کارشناسی، کارشناسی ارشد و دکتری سایر رشته‌های تحصیلی و به خصوص متخصصین علوم کشاورزی و دامی باشد.

همچنین، پیشرفت‌های زیادی در توسعه مدل‌های آماری و کاربرد آن‌ها در علوم مختلف صورت گرفته است. یکی از مسائل مورد اهمیت کاربرد مدل‌های آماری در علوم دامی می‌باشد که تاکنون کتاب‌های متعددی در این زمینه نوشته شده و توسط پژوهشگران و تخصصین این رشته نیز به زبان فارسی ترجمه شده‌اند. اما تقریباً تمامی این کتاب‌ها حل مسائل را با نرم افزار آماری SAS ارائه داده‌اند که به دلیل عدم دسترس بودن این نرم افزار و رجحان نرم افزار R از هر جهت، ضرورت تهیه این کتاب وجود داشت. از این رو برآن شدیم که این کتاب نوشته و گردآوری گردد.

این کتاب در دو قسمت کلی با نام‌های روش‌های پیشرفته آماری و مدل‌های خطی در علوم دامی تهیه شده و شامل ۱۳ فصل است که در فصل اول نحوه استفاده و آشنایی با نرم افزار R بیان شده است.

در فصل دوم برخی روش‌های آماری توصیفی و استنباطی مقدماتی بیان و با ذکر مثال‌هایی با نرم افزار R تشریح شده‌اند. فصل سوم به جبر خطی و محاسبات ماتریسی و نحوه انجام آن‌ها در نرم افزار R اختصاص یافته است. در فصل چهارم برخی از روش‌های آمار استنباطی شامل برآورد فاصله‌ای پارامتری و آزمون فرضیه‌های آماری پارامتری بیان و با نرم افزار آماری R نحوه حل مسائل بیان شده است. الگوی رگرسیون‌های خطی و غیرخطی و مدل‌های رگرسیون خطی چندگانه با استفاده از نرم افزار R در فصل پنجم تشریح شده است. در فصل ششم به بررسی انواع رگرسیون‌های با پاسخ‌های رشته‌ای (کیفی یا گسسته) پرداخته شده است و مثال‌های متعددی با نرم افزار R ذکر شده است. در فصل هفتم تحلیل بیشتری از برازش مدل‌های رگرسیونی شامل پذیره‌های زیر بنایی و چگونگی بررسی آن‌ها، نحوه گزینش متغیرهای توضیحی، نقاط دور افتاده، موثر و... با استفاده از نرم افزار R بیان شده است.

در قسمت دوم که از فصل هشتم آغاز می‌شود، انواع مدل‌های خطی در علوم دامی بررسی شده است. بنابراین، در فصل هشتم بهترین پیش‌بینی ناریب خطی ارزش‌های اصلاحی و مدل دام معرفی و تشریح شده است. در فصل نهم به بررسی برآورد اثرات ژنتیکی مادری در مدل دام و حل مسائل مرتبط با نرم افزار R پرداخته شده است. برآورد اثر ژنتیکی و محیطی مادری با استفاده از مدل دام و حل چند مثال با استفاده از نرم افزار R در فصل دهم بیان شده است. فصل یازدهم به معرفی برآورد اثرات ژنتیکی مدل دام چند صفتی حل مثال‌های کاربردی اختصاص دارد. در فصل دوازدهم برآورد اثرات ژنتیکی غیر افزایشی با استفاده از نرم افزار R ذکر شده است. رگرسیون تصادفی در فصل سیزدهم بیان و مثال‌هایی با نرم افزار R توضیح داده شده است. در پیوست‌های (الف) تا (ج) داده‌های استفاده شده در فصول قبل آمده است.

در پایان بر خود لازم می‌دانیم تا از زحمات دانشجویان مقطع دکتری علوم دامی دانشگاه فردوسی مشهد، آقایان دکتر مسلم شجاعی و دکتر صابر جلوخانی نیارکی که در تهیه برخی از مطالب اولیه همکاری داشتند، تشکر و قدردانی کنیم.

از تمامی همکاران گروه آمار دانشگاه فردوسی مشهد، به خصوص استاد گرامی جناب آقای دکتر صادق‌پور گیله‌ده که زحمت ویراستاری علمی کتاب را برعهده داشتند، کمال تشکر و قدردانی را دارد. همچنین، از معاونت محترم پژوهشی و انتشارات دانشگاه فردوسی مشهد به جهت بررسی و اجازه چاپ و نشر این کتاب تشکر و قدردانی می‌شود.

با احترام

مهدی جبباری نوقابی

عضو هیئت علمی گروه آمار دانشگاه فردوسی مشهد

تابستان ۱۳۹۵

فصل اول

مقدمه‌ای بر نرم افزار آماری R

۱-۱ - نرم افزار آماری R

با پیشرفت علوم و گسترش فناوری، اهمیت استفاده از روش‌های آماری در علوم مختلف بیش از پیش مورد توجه قرار گرفته است و فراگیری آمار کاربردی در هر رشته جزء ملزومات آن رشته محسوب می‌گردد. فرآیند آنالیز آماری کمک می‌کند تا پژوهشگر بتواند از داده‌های اولیه، اطلاعات مورد نیاز خود را استخراج کند و در صورت لزوم نتایج را تعمیم دهد. اگر حجم داده‌ها بزرگ باشد، استفاده از روش‌های مختلف آنالیز آماری بسیار خسته کننده و مشکل خواهد بود. امروزه انواع نرم افزارهای آماری موجود، می‌توانند آنالیزهای آماری مختلفی را انجام دهند. نرم افزار R یکی از برنامه‌های کاربردی در زمینه تجزیه و تحلیل‌های آماری می‌باشد. این نرم افزار آماری با قابلیت انجام توصیفی زیبا و گویا از اطلاعات، شامل رسم نمودارها و شکل‌های گوناگون و محاسبات مربوط به شاخص‌های آمار توصیفی تمایل مرکزی، پراکندگی و پراکندگی نسبی شامل میانگین، انحراف معیار، واریانس، میانه و غیره می‌باشد.

نرم افزار R یک زبان برنامه‌نویسی و محیط نرم‌افزاری برای محاسبات آماری و تحلیل داده می‌باشد، که بر اساس زبان‌های اس و اسکیم پیاده‌سازی شده است. پروژه R از سال ۱۹۹۵ میلادی

در گروه آمار دانشگاه اوکلند^۱ توسط آقایان روبرت جنتلمن^۲ و راس ایاکا^۳ (به همین علت نام R برای آن انتخاب گردید) آغاز شد و تا زمان حال مخاطبین فراوانی یافت. در حال حاضر این زبان را یک گروه بین‌المللی نگهداری می‌کند و داوطلبانه به توسعه آن می‌پردازد.

نرم افزار R، حاوی محدوده گسترده‌ای از روش‌های آماری (از جمله: مدل‌سازی خطی و غیرخطی؛ آزمون‌های کلاسیک آماری؛ تحلیل سری‌های زمانی؛ رده‌بندی؛ خوشه‌بندی و ...) و قابلیت‌های گرافیکی است.

گرچه نرم‌افزار R اغلب به منظور انجام محاسبات آماری به کار می‌رود، اما از قابلیت به کارگیری در محاسبات ماتریسی نیز برخوردار می‌باشد و در این زمینه، هم پای نرم‌افزارهایی چون اکتاو و نسخه تجاری آن متلب^۴ می‌باشد. نرم افزار R دارای محیط خط فرمان برای ورود و اجرای دستورات است.

۱-۲- نحوه نصب نرم افزار R

ابتدا فایل اجرایی R-3.2.3 for Windows را از درگاه نرم افزار R ذخیره نموده و با دوبر کلیک متوالی روی آن نصب می‌شود. پس از پایان عمل نصب روی desktop رایانه شما یک آیکون به شکل حرف R قرار می‌گیرد. اگر روی حرف R دو بار کلیک کنید، صفحه‌ای باز می‌شود که R Console نام دارد. در این صفحه توضیحات کوتاهی وجود دارد و پس از آن علامت «>» ملاحظه می‌شود که در مقابل آن می‌توان عملیات مورد نظر را انجام داد.

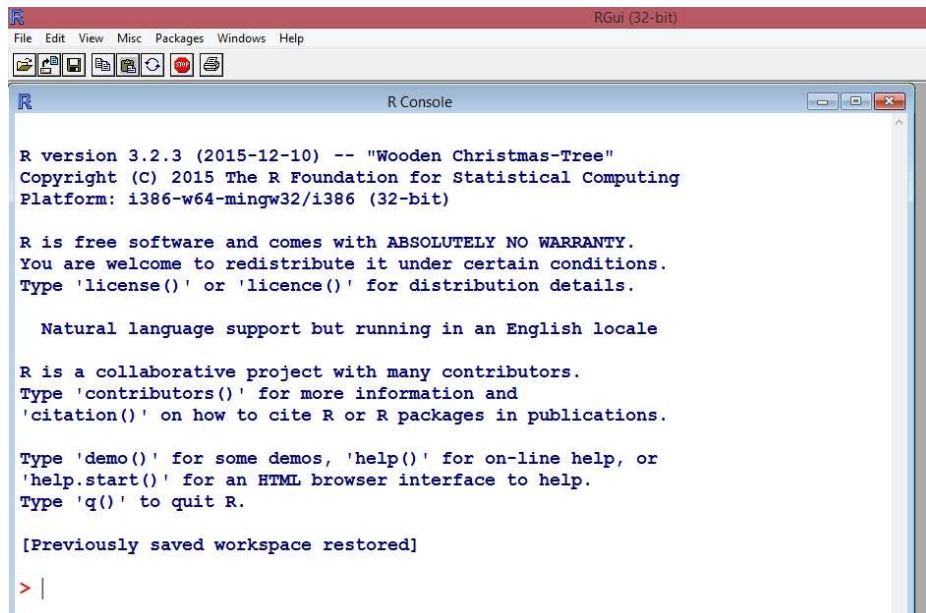
۱-۳- مزایای نرم افزار R

برخی از مزایای نرم افزار R به صورت زیر است:

- قابلیت دسترسی رایگان آن برای همه؛
- قابلیت اجرا در بیشتر سیستم عامل‌ها (ویندوز، لینوکس و مکینتاش)؛
- قابلیت توسعه بی‌نظیری همچون دیگر نرم افزارهای متن باز؛
- امکان دسترسی رایگان به بسته‌های نرم افزاری از طریق درگاه سازنده؛

1. Auckland
2. Robert Gentleman
3. Ross Ihaka
4. Matlab

- به روز بودن بسته‌های نرم‌افزاری مربوط به نرم افزار R؛
- فراهم ساختن بسته‌های نرم‌افزاری و توابع جدید برای همه؛
- امکان برنامه نویسی (نوشتن کد و تابع)؛
- مشترک بودن زبان برنامه نویسی آن با S-PLUS.



```
RGui (32-bit)
File Edit View Misc Packages Windows Help
R Console
R version 3.2.3 (2015-12-10) -- "Wooden Christmas-Tree"
Copyright (C) 2015 The R Foundation for Statistical Computing
Platform: i386-w64-mingw32/i386 (32-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

> |
```

شکل ۱-۱: صفحه R console.

۱-۴- محدودیت‌های نرم افزار R

- برخی محدودیت‌های نرم افزار R به شرح زیر است:
- برای استفاده و برنامه نویسی با آن، باید فرامین آن را آموخت؛
 - زبان نرم افزار R دارای مفسر می‌باشد، به این معنی که برای اجرا ترجمه نمی‌شود و به صورت فایل اجرایی (*.exe) در نمی‌آید.

۱-۵- آشنایی با منوهای نرم افزار R

نوار منو در نرم افزار R بسیار شبیه به برنامه‌های مبتنی بر ویندوز می‌باشد و شامل منوهای دریافت

است که به اختصار در ادامه توضیح داده شده است. بسیاری از این قابلیت‌های ارائه شده توسط منوها، با استفاده از دستورات استاندارد ویندوز فراهم می‌باشد (برای مثال، CTRL+C).

۱-۵-۱- File منوی

New script & Open script: باز کردن سند جدید و یا سند ذخیره شده پیشی برای وارد کردن دستورات می‌باشد.

Display file(s): نمایش فایل‌های قابل شناسایی در نرم افزار.

Load Workspace & Save Workspace: فراخوانی و ذخیره صفحه کاری.

Load History & Save History: فراخوانی و ذخیره تاریخچه.

Change dir: تغییر مسیر (دایرکتوری) کاری.

Save to File: ذخیره دستورات و خروجی به صورت فایل متنی می‌باشد.

۱-۵-۲- Edit منوی

Paste commands only: فقط چسباندن دستورات.

Clear console: پاک کردن صفحه دستورات.

Data Editor: ویرایش داده‌ها.

GUI Preferences: تنظیم پیش فرض‌ها می‌باشد.

۱-۵-۳- Misc منوی

Stop current computation: توقف محاسبه فعلی یا کل محاسبات.

List object: نمایش اشیاء R.

Remove all objects: حذف تمام اشیاء.

List search path: نمایش بسته‌های در حال استفاده.

۱-۵-۴- Packages منوی

Load package: فراخوانی بسته.

Set CRAN mirror: تنظیم سرور پیش فرض (برای مثال، سرور ایران واقع در دانشگاه فردوسی مشهد).

Install package(s): نصب برخط (آن‌لاین) بسته‌ها.

Update packages: بروز رسانی بسته‌ها می‌باشد.

۱-۴-۵-۱- نصب و فراخوانی بسته‌ها

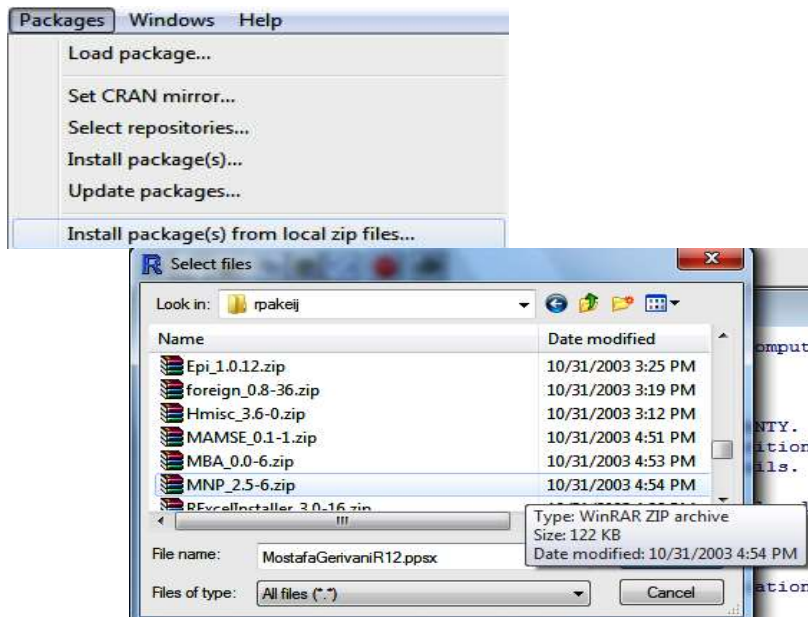
در هنگام کار با نرم افزار R، گاهی اوقات برای دسترسی به برخی از توابع خاص، نیازمند استفاده از بسته‌هایی^۱ هستیم. برای نصب این بسته‌ها به چند روش می‌توان عمل کرد:

الف) ابتدا بسته مورد نظر را از حالت فشرده خارج ساخته و سپس آن را در شاخه `C:\Program Files\R\R 3.2.3\library` کپی کرده و سپس در محیط نرم افزار دستور `library` یا `require` را اجرا می‌کنیم.

> library(نام بسته)

ب) در این روش نیازی به خارج ساختن بسته از حالت فشرده نمی‌باشد. ابتدا از منوی `packages` گزینه آخر (`Install package from...`) را انتخاب کرده و سپس از منوی باز شده، بسته‌های مورد نظر را انتخاب کنید.

حال، دوباره از منوی `package` گزینه `load package` را انتخاب کرده و در منوی باز شده، بسته مورد نظر را انتخاب یا از همان دستور `library` استفاده می‌کنیم.

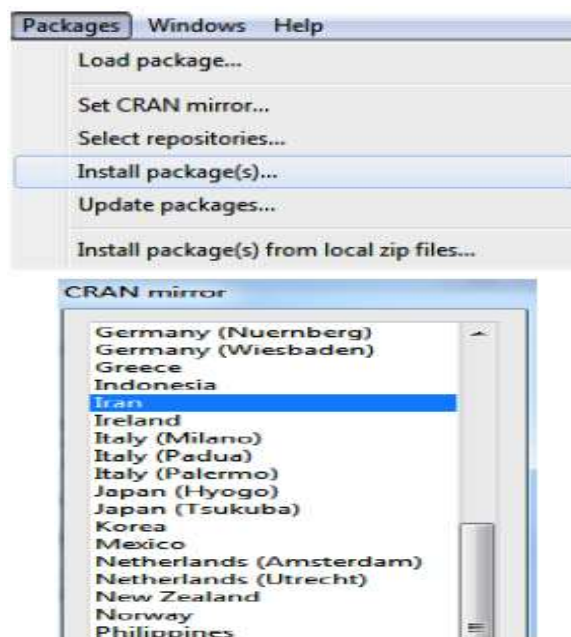


شکل ۱-۲: چگونگی نصب و به کارگیری (Load) یک بسته در نرم افزار R

ج) در این روش، نرم افزار به گونه مستقیم به شبکه ارتباط جمعی (اینترنت) وصل شده و بسته مورد نظر را ذخیره می‌کند. لذا برای استفاده از این روش، دسترسی به شبکه ارتباط جمعی لازم است.

`install.packages('نام بسته', rep="http://cran.um.ac.ir")`

این کار را می‌توان از طریق منویی به صورت زیر انجام داد. ابتدا از منوی package گزینه `install.package` را انتخاب کنید. در صفحه باز شده نام ایران را انتخاب و ok را انتخاب کنید.



شکل ۱-۳: نحوه نصب یک بسته در نرم افزار R با استفاده از شبکه ارتباط جمعی

۱-۵-۵-۵- Help منوی

Console: راهنمای کلیدهای میان‌بر در صفحه Console.

FAQ on R: بیشترین سؤالات کاربران.

Manuals: چند فایل راهنما به صورت PDF.

R functions: راهنمای توابع نرم افزار R.

Html help: راهنمای کامل نرم افزار.