

دانشگاه پیام نور
۱۰۴۶

گروه آمار
(۱/۲۵)

بهبودیان، جواد، ۱۳۱۰ -
رگرسیون (رشته آمار) / تألیف جواد بهبودیان. — تهران: دانشگاه پیام نور، ۱۳۸۳.
پازده، ۳۰۵ ص. : جدول، نمودار. — (دانشگاه پیام نور؛ ۱۰۴۶. گروه آمار؛
(۱/۲۵)
فهرست نویسی بر اساس اطلاعات فیبا.
واژه نامه.
کتابنامه.
چاپ پنجم.
۱. آموزش از راه دور - ایران. ۲. برگشت (ریاضیات). الف. دانشگاه پیام نور.
ب. عنوان.
۳۷۸ / ۱۷۵۰۹۵۵ LC ۵۸۰۸ / الف ۹ ب ۹۱۴
۸۳ - ۸۹۱ م کتابخانه ملی ایران

۱۹۱ دوی

کتابخانه مرکزی پژوهش های آموزشی و تربیتی
شماره ثبت: ۱۴۳۴
تاریخ: ۱۴۰۴/۰۴/۲۲



دانشگاه پیام نور
رگرسیون

دکتر جواد بهبودیان

ویراستار علمی: دکتر نرگس عباسی

حروفچینی و نمونه خوانی: مدیریت تدوین

رسم، طراح جلد و صفحه آرا: فرشته فلاح دوست

لیتوگرافی، چاپ و صحافی: انتشارات دانشگاه پیام نور

شمارگان: ۶۰۰۰ نسخه

نوبت و تاریخ چاپ: چاپ پنجم در سال آذر ۱۳۸۲، چاپ دوم آزمایشی مرداد ۱۳۸۵

شابک ۹ - ۰۶۶ - ۳۸۷ - ۹۶۴

ISBN 964 - 387 - 066 - 9

کلیه حقوق برای دانشگاه پیام نور محفوظ است.

قیمت: ۱۸۱۰۰ ریال

بسم الله الرحمن الرحيم

پیشگفتار ناشر

کتابهای دانشگاه پیام نور حسب مورد و با توجه به شرایط مختلف به صورت درسنامه، آزمایشی، قطعی، متون آزمایشگاهی، فرادرسی، و کمک‌درسی چاپ می‌شود. کتاب درسنامه (د) نخستین ثمره کوششهای علمی صاحب اثر است که براساس نیازهای درسی دانشجویان و سرفصلهای مصوب تهیه می‌شود و پس از داوری علمی در گروههای آموزشی، بدون طراحی آموزشی و ویرایش چاپ می‌شود. با تجدیدنظر صاحب اثر و دریافت بازخوردها و اصلاح نارساییها، درسنامه با طراحی آموزشی، ویرایش، و طراحی فنی - هنری به صورت آزمایشی (آ) چاپ می‌شود. با دریافت نظرهای اصلاحی، صاحب اثر در کتاب تجدید نظر می‌کند و کتاب به صورت قطعی (ق) چاپ می‌شود. در صورت ضرورت، در کتابهای چاپ قطعی نیز تجدید نظرهای اساسی به عمل می‌آید یا متناسب با پیشرفت علوم و فناوری بازنویسی می‌شوند. متون آزمایشگاهی (م) متونی است که دانشجویان با استفاده از آن و راهنمایی مربیان کارهای عملی آزمایشگاهی را انجام می‌دهند. کتابهای فرادرسی (ف) و کمک‌درسی (ک) به منظور غنی‌تر کردن منابع درسی دانشگاهی تهیه می‌شوند. کتابهای فرادرسی با تأیید معاونت پژوهشی و کتابهای کمک‌درسی با تأیید شورای انتشارات تهیه می‌شوند.

مدیریت تدوین

فهرست

| نه یازده | | پیشگفتار | تاریخچه‌های کوتاه | |
|-------------|--|----------|-------------------|--|
| ۱ | | | | ۱- مفهوم و ویژگیهای همبستگی و وابستگی |
| ۲ | | | | ۱-۱ امید ریاضی تابعی از چند متغیر تصادفی |
| ۱۱ | | | | ۱-۲ کواریانس و همبستگی |
| ۱۳ | | | | ۱-۲-۱ ضریب همبستگی دو متغیر تصادفی |
| ۱۷ | | | | ۱-۳ همبستگی و وابستگی |
| ۲۱ | | | | ۱-۴ برآوردیابی ضریب همبستگی |
| ۲۵ | | | | ۱-۵ توزیع آماره‌ی R |
| ۳۰ | | | | ۱-۶ راهنمایی تمرینهای فصل ۱ |
| ۴۷ | | | | ۲- تابع پیش‌بینی‌کننده و خط رگرسیون |
| ۴۷ | | | | ۲-۱ بهترین تابع پیش‌بینی‌کننده |
| ۵۰ | | | | ۲-۲ بهترین تابع پیش‌بینی‌کننده خطی |
| ۵۳ | | | | ۲-۳ برآورد بهترین تابع پیش‌بینی‌کننده خطی |
| ۵۶ | | | | ۲-۴ راهنمایی تمرینهای فصل ۲ |
| ۶۱ | | | | ۳- رگرسیون خطی ساده |
| ۶۱ | | | | ۳-۱ مدل خطی ساده |
| ۶۳ | | | | ۳-۲ برآورد پارامترهای مدل خطی ساده با روش کمترین مجموع توانهای دوم |
| ۶۸ | | | | ۳-۳ تعبیر هندسی روش LS |
| ۶۹ | | | | ۳-۴ توزیع \hat{a} ، \hat{b} و $\hat{\sigma}^2$ با فرض خطای نرمال |
| ۷۲ | | | | ۳-۵ فاصله اطمینان و آزمون فرض آماری |
| ۷۵ | | | | ۳-۶ انحراف یک متغیر وابسته از معدل این متغیرها |
| ۷۷ | | | | ۳-۷ آنالیز واریانس برای یک مدل خطی ساده |
| ۸۲ | | | | ۳-۸ ضریب تعیین |

تستی و تکرینی

۸۴
۸۷
۸۹
۹۲
۱۰۵
۱۰۷
۱۱۳
۱۱۳
۱۲۷
۱۴۲
۱۵۰
۱۵۴
۱۵۷
۱۶۷
۱۶۷
۱۷۴
۱۸۲
۱۸۹
۱۸۹
۱۹۸
۲۰۱
۲۰۷
۲۱۰
۲۱۶
۲۲۱
۲۲۱
۲۳۶
۲۴۵
۲۴۹
۲۴۹
۲۵۳
۲۵۸
۲۶۱
۲۶۵
۲۶۹
۲۷۲
۲۷۹

۳-۹ مدل‌های غیرخطی
۳-۱۰ مدل خطی وقتی که هر دو متغیر دستخوش خطا باشند
۳-۱۱ مدل خطی ساده گذرنده از مبدأ
۳-۱۲ چند نکته درباره مدل خطی ساده
۳-۱۳ یک مثال کاربردی
۳-۱۴ راهنمایی تمرینهای فصل سوم

۴- فضای برداری با ضرب داخلی و تصویر
۴-۱ فضای برداری با ضرب داخلی
۴-۲ تصویر یک بردار
۴-۳ ماتریس تصویر
۴-۴ ماتریس همیشه مثبت
۴-۵ وارون گسترده یک ماتریس
۴-۶ راهنمایی تمرینهای فصل ۴

دانش

۵- بردار تصادفی و بردار نرمال
۵-۱ بردار تصادفی و ویژگیهای آن
۵-۲ بردار تصادفی نرمال
۵-۳ راهنمایی تمرینهای فصل ۵

تیم ریاضی

۶- برآورد پارامترها در مدل‌های خطی چندمتغیری
۶-۱ مدل‌های خطی چندمتغیری
۶-۲ تعبیر هندسی مدل‌های خطی چندمتغیری
۶-۳ مختصات کانونی
۶-۴ مدل خطی تحت فرض خطای نرمال
۶-۵ چند نکته درباره مدل خطی
۶-۶ راهنمایی تمرینهای فصل ۶

بسیار

۷- استنباط آماری در مدل‌های خطی چندمتغیری
۷-۱ آزمون آماری فرضهای خطی در مدل (CN)
۷-۲ چند مثال درباره مدل‌های خطی
۷-۳ راهنمایی تمرینهای فصل ۷

بسیار

۸- موضوعهای گوناگون در رگرسیون
۸-۱ ضریب همبستگی جزئی و چندگانه
۸-۲ قضیه کاکران و کاربرد آن در رگرسیون
۸-۳ رگرسیون با برخی متغیرهای مستقل صفر و یک
۸-۴ همخطی و رگرسیون برآمده
۸-۵ رگرسیون لوجستیک چیست؟
۸-۶ رگرسیون با کمترین مجموع قدرمطلق خطا
۸-۷ راهنمایی تمرینهای فصل ۸

الف. متغیرهای تصادفی X^2 نامرکزی، F نامرکزی و T نامرکزی

شش

| | |
|-----|--------------------------------------|
| ۲۷۹ | ۱. الف متغیر تصادفی χ^2 نامرکزی |
| ۲۸۰ | ۲. الف متغیر تصادفی F نامرکزی |
| ۲۸۰ | ۳. الف متغیر تصادفی T نامرکزی |
| ۲۸۱ | ب. P- مقدار |
| ۲۸۱ | ۱- ب طرز محاسبه P- مقدار |
| ۲۸۵ | پ. روش تکراری نیوتن |
| ۲۸۹ | جداول آماری |
| ۲۹۹ | واژه‌نامه فارسی - انگلیسی |
| ۳۰۱ | واژه‌نامه انگلیسی - فارسی |
| ۳۰۳ | فهرست راهنما |
| ۳۰۵ | مراجع |

پیشگفتار

آنهایی که در زمینه علوم زیستی یا تجربی یا اجتماعی مطالعه و پژوهش می‌کنند معمولاً با دو نوع متغیر روبه‌رو می‌شوند. یک نوع آن اغلب قابل کنترل و نوع دیگر تصادفی می‌باشد. مثلاً سن یک نوزاد (برحسب روز) متغیر قابل کنترل ولی وزن او (برحسب گرم) متغیر تصادفی است. بدیهی است که نوزادهای ۱۰ روزه دارای وزنه‌های گوناگون می‌باشند. کشف ارتباط و مدل‌سازی میان این‌گونه متغیرها، به‌منظور پیش‌بینی مقدار متغیر تصادفی و مطالعه‌ی عوامل مؤثر در مقدار آن، برای پژوهش‌های علمی ضرورت دارد.

ضریب همبستگی و رگرسیون (یا برگشت) گفتاریست آماری درباره‌ی این‌گونه متغیرها. این گفتار به سبب تئوری پربار آن، از نظر آمار ریاضی، و کاربرد گسترده آن در رشته‌های گوناگون دارای ارزش آماری می‌باشد. در عصر ما ده‌ها کتاب و صدها مقاله موضوع را با روش‌های جالب بررسی کرده‌اند.

این کتاب برای آشنایی با ضریب همبستگی و رگرسیون می‌باشد که از نظر تاریخی هر دو موضوع با هم توسط گالتن، دانشمند انگلیسی، در نیمه دوم قرن نوزدهم ارائه شده است. آنهایی که با مقدمات آمار ریاضی آشنا هستند می‌توانند آن را به‌عنوان کتاب درسی برای دوره‌های کارشناسی آمار یا برخی رشته‌های دیگر به مدت یک

نیمسال به کار برند. پژوهشگران هم می‌توانند در کارهای پژوهشی خود از این کتاب استفاده کنند. ناگفته نماند که چنین کتابی نمی‌تواند شامل تمام جوانب موضوع و کاربردهای بی‌شمار آن باشد. ولی نویسنده سعی کرده است مطالب مختلف را به کوتاهی، با زبان ساده‌ای از آمار ریاضی، با مثالهای کاربردی و تمرینهای آموزنده بنگارد. تمام تمرینها دارای پاسخ و راهنمایی کوتاه می‌باشند. از ذکر برنامه‌های کامپیوتری، برای جلوگیری از حجم کتاب و برای اینکه خواننده بتواند نرم‌افزار دلخواه خود را برگزیند، دوری شده است. این کتاب براساس تجربه شخصی نویسنده همراه با پاسخ و راهنمایی تمرینها تهیه شده است.

از همه آنهایی که به نحوی برای آمادگی این کتاب یاری کرده‌اند صمیمانه سپاسگزارم. به‌ویژه همکاری ارزنده آقای امین قلمفرسای مستوفی، کارشناس آمار که فصل به فصل کتاب را با دقت خوانده، برای بهبود آن پیشنهادهای مفید داده، با حوصله غلطگیری کرده، و پاسخ و راهنمایی تمرینها را تهیه کرده‌اند مورد کمال امتنان و قدردانی می‌باشد. سرانجام از خانم بهناز میرزایی، منشی بخش آمار دانشکده علوم دانشگاه شیراز، به‌خاطر حوصله فراوان برای ماشین‌نویسی کامپیوتری کتاب بی‌اندازه ممنونم.

جواد بهبودیان

بخش آمار دانشکده علوم دانشگاه شیراز

شهریور ۱۳۸۲

تاریخچه‌ای کوتاه

بسیاری می‌پندارند که کارل پیرسن (Karl Pearson 1857-1936) آماردان معروف انگلیسی ضریب همبستگی را معرفی کرده است. البته او از نظر ریاضی درباره این معیار کار ریاضی ارزنده‌ای انجام داد و از این‌رو این معیار را ضریب همبستگی گشتاوری پیرسن می‌نامند. ولی از نظر تاریخی فرانسیس گالتن (Francis Galton 1822-1911) دانشمند انگلیسی نخستین کسی است که واژه همبستگی (Correlation) را در کتاب وراثت خود به‌کار برده است.

گالتن، پسر عموی چارلز داروین (Charles Darwin)، نخستین دانشمند قرن نوزدهم است که برای مطالعه مسائل انسانی و حیاتی مفهوم جدید همبستگی و رگرسیون را برای کارهای آماری در مورد وزن دانه‌های نخود سبز و نیز در مورد اندازه قامت پدر و پسر به‌کار برد. با اینکه نمی‌توان او را ریاضیدان پنداشت ولی صاحب اندیشه آماری بود و به آمار عشق می‌ورزید. او می‌گوید «برخی از آمار بدشان می‌آید ولی من آمار را پر از زیبایی و کشش می‌دانم».

کارل پیرسن، همکار گالتن و یکی از پژوهشگران در آزمایشگاه او، شرح حال گالتن را پس از مرگش در ۱۹۱۱ در ۴ جلد منتشر کرد. گالتن نه تنها رگرسیون ساده را معرفی کرد، بلکه در کارهای عملی خود هسته اساسی رگرسیون چندگانه را هم بررسی