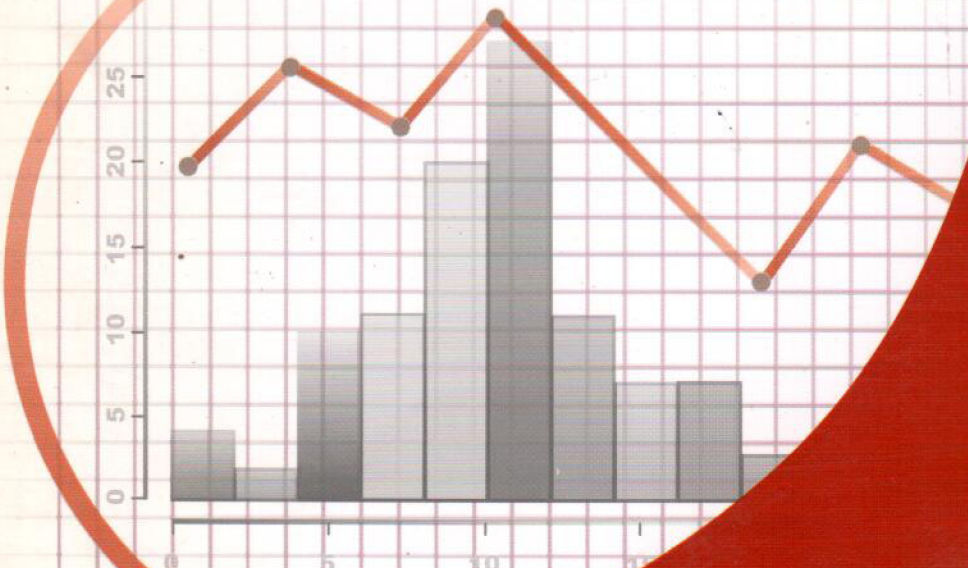
دانشگاه علوم بهزیستی و توانبخشی

آمار تحلیلی

بامعرفی نرم افزار R



مؤلفان:

دکتر غلامرضا جندقی (دانشیار دانشگاه تهران)

علی حمیدی زاده، دکتر مسعود کریملو

بسم الله الرحمن الرحیم

آمار تحلیلی
با معرفی نرم افزار R

نویسندگان:دکتر غلامرضا جندقی-علی حمیدی زاده-دکتر مسعود کریملو



سرشناسه: جندقی، غلامرضا، ۱۳۳۹-
عنوان و نام پدیدآور: آمار تحلیلی با معرفی نرم افزارR/ نویسندگان غلامرضا جندقی، علی حمیدی زاده، مسعود کریملو.
مشخصات نشر: تهران: دانشگاه علوم بهزیستی و توانبخشی،۱۳۸۸.
مشخصات ظاهری:۲۹۲ص.:مصور، جدول، نمودار.
شابک:۴۲۰۰۰ریال:978-964-8863-43-7
وضعیت فهرست نویسی: فیپا
موضوع: آمار ریاضی- - داده پردازی
شناسه افزوده: حمیدی زاده،علی،۱۳۶۱-
شناسه افزوده: کریملو، مسعود، ۱۳۴۰-
شناسه افزوده: دانشگاه علوم بهزیستی و توانبخشی
رده بندی کنگره:۱۳۸۸ ۸آ۹ج/۲۷۶/٤ QA
رده بندی دیویی: ۵۱۹/۵۰۲۸۵
شماره کتابشناسی ملی: ۱۹۰٤۸۱۱

دانشگاه علوم بهزیستی و توانبخشی

نام کتاب:آمار تحلیلی با معرفی نرم افزار R
نویسندگان: دکتر غلامرضا جندقی- علی حمیدی زاده- دکتر مسعود کریملو
ویراستار: دکتر کمال اعظم
طراح جلد: فرزانه فلاح زاده
شابک :۹۷۸-۹۶۴-۸۸۶۳-۴۳-۷
قیمت :۴۲۰۰۰ریال
سال انتشار: ۱۳۸۸
چاپ: آرویج
ناشر: دانشگاه علوم بهزیستی و توانبخشی
نشانی :
اوین، بلوار دانشجو، خیابان کودکیار، دانشگاه علوم بهزیستی و توانبخشی
انتشارات دانشگاه تلفکس: ۲۲۱۸۰۰۸۱

همه حقوق برای ناشر محفوظ است.

پیشگفتار

کتاب حاضر حاصل سالها تدریس درس آمار پیشرفته در دوره‌های کارشناسی ارشد می باشد. این کتاب با هدف پوشش سرفصل درس تحلیل آماری برای دوره‌های کارشناسی ارشد رشته‌های علوم انسانی تألیف شده است. یکی از ویژگی‌های منحصر بفرد این کتاب، معرفی نرم افزار R در تجزیه و تحلیل داده‌های آماری است. ساختار کتاب مشتمل بر دو قسمت می‌باشد. در قسمت اول به مباحث آمار تحلیلی پرداخته شده است. بدین منظور چهار فصل در نظر گرفته شده است. در فصل اول، نظریه برآورد مطرح شده است. در فصل دوم آزمون فرضیه و روش‌های مختلف آمار پارامتری شامل، آزمون‌های پارامتری یک جامعه، دو جامعه و چند جامعه آورده شده است. در فصل سوم به مباحث رگرسیون خطی ساده و چند گانه پرداخته و نهایتاً فصل چهارم نیز، به آزمون‌های ناپارامتری شایع در تحقیقات مدیریت، اختصاص دارد.

در قسمت دوم کتاب ضمن معرفی نرم‌افزار R، چگونگی توصیف و تحلیل داده‌های آماری در این نرم افزار به طور مشروح توضیح داده شده است. این بخش از کتاب همچون بخش قبلی، در بردارنده چهار فصل می‌باشد. فصل اول به مقدمات R اشاره دارد، در فصل دوم آمار توصیفی مطرح شده و در فصل سوم نیز مباحث مرتبط با احتمال، نمونه‌گیری و برآورد مورد بررسی قرار می‌گیرد. فصل چهارم با عنوان آمار تحلیلی است که مشتمل بر آزمون‌های پارامتری و ناپارامتری می‌باشد.

مؤلفین امیدوارند از نظرات اساتید و دانشجویان و محققان در جهت بهبود این اثر در چاپ‌های بعدی استفاده نمایند. لذا خواهشمندیم بر مؤلفین منت نهاده و نظرات

اصلاحی و پیشنهادهای خود را به پست الکترونیکی jandaghi@ut.ac.ir ویا mkarimlo@uswr.ac.ir ارسال نمایید.

در خاتمه از آقایان **محمدجواد دهقان اشکذری** و **سید مجتبی حسینی فرد** که کار ویراستاری و صفحه‌آرایی و خانم **فرزانه فلاح‌زاده** که طراحی جلد را متقبل شده‌اند، کمال تشکر را داشته و از یکایک این عزیزان قدردانی می‌نماییم.

دکتر غلامرضا جندقی

علی حمیدی زاده

دکتر مسعود کریملو

خرداد ۸۸

فهرست مطالب

قسمت اول: آمار تحلیلی

فصل اول: برآورد ۱۳

مقدمه ۱٤

تعریف برآورد ۱٤

خواص برآورد کننده‌ها ۱٤

تعریف ۱٦

انواع برآورد ۱۷

برآورد نقطه‌ای ۱۷

برآورد فاصله‌ای ۱۸

فاصله اطمینان برای میانگین جامعه ۱۸

فاصله اطمینان برای μ وقتی σ^2 معلوم باشد (جامعه نرمال) ۲۰

فرمول حجم نمونه برای برآورد μ ۲۲

فاصله اطمینان برای μ وقتی σ^2 مجهول باشد(جامعه نرمال) ۲۳

فاصله اطمینان برای واریانس یک جامعه ۲۵

فاصله اطمینان برای نسبت یک صفت در جامعه ۲٦

فرمول حجم نمونه برای برآورد یک نسبت ۲۷

برآورد فاصله‌ای تفاضل میانگین دو جامعه $(\mu_1 - \mu_2)$ ۲۸

برآورد فاصله‌ای تفاضل میانگین دو جامعه (واریانس‌ها مجهول و مساوی) ۳۰

برآورد فاصله‌ای تفاضل میانگین دو جامعه (واریانس‌ها مجهول و نامساوی) ۳۱

برآورد فاصله‌ای تفاضل نسبت موفقیت در دو جامعه ۳۲

برآورد فاصله‌ای نسبت واریانس دو جامعهٔ آماری ۳۳

تمرینات ۳٤

فصل دوم: آزمون فرضیه ۳۷

مقدمه ۳۸

فرضیه ۳۸
آزمون فرضیه ۳۸
فرضیه صفر ۳۹
فرضیه مقابل ۳۹
آماره آزمون ۳۹
ناحیه بحرانی (رد) ۳۹
خطای نوع اول (سطح معنی دار بودن، اندازه آزمون، اندازه ناحیه) ٤۰
خطای نوع دوم ٤۰
توان آزمون (قدرت آزمون، قدرت براندازی فرضیه صفر) ٤۲
آزمون فرضیه دو دامنه ٤۳
آزمون فرضیه یک دامنه ٤۳
مراحل آزمون فرضیه ٤٤
آزمون میانگین یک جامعه ٤٥
حالت اول: جامعه مورد نمونه گیری نرمال و واریانس جامعه معلوم است. ٤٥
حالت دوم: جامعه مورد نمونه گیری نرمال و واریانس جامعه مجهول است. ٤۷
آزمون اختلاف میانگین‌های دو جامعه مستقل ٤۹
آزمون اختلاف میانگین‌ها در دو جامعه وابسته ٥٤
آزمون فرضیه نسبت موفقیت در جامعه (P) ٥۷
آزمون تفاوت بین دو نسبت ٦۰
آزمون واریانس جامعه ٦۳
آزمون واریانس دو جامعه ٦٦
مقایسه میانگین بیش از دو جامعه (آنالیز واریانس) ۷۰
آنالیز واریانس یک عاملی (طرح کاملاً تصادفی) ۷۱
آزمون‌های تعقیبی ۷٥
آنالیز واریانس دو عاملی (طرح بلوک تصادفی) ۷۷
آزمون اثر تیمار ۷۹
آزمون اثر بلوک ۷۹

آنالیز واریانس دو عاملی با تکرار ۸۳
بررسی اثر عامل اول (عامل سطری) ۸۶
بررسی اثر عامل دوم (عامل ستونی) ۸۶
بررسی اثر متقابل دو عامل ۸۶
تمرینات ۸۹
فصل سوم: رگرسیون ۹۵
مقدمه ۹۶
مدل رگرسیون خطی ساده ۹۶
برازش مدل رگرسیون خطی به داده‌ها ۹۷
ضریب تعیین مدل ۱۰۰
استنباط در تحلیل رگرسیونی ۱۰۲
استنباط در مورد شیب خط ۱۰۲
استنباط در مورد عرض از مبدأ ۱۰۳
استنباط در مورد مقادیر پیش بینی شده ۱۰۵
استنباط در مورد میانگین یک مقدار پاسخ جدید ۱۰۷
آزمون خطی بودن رگرسیون ۱۰۸
رگرسیون چندگانه ۱۰۹
بیان ماتریسی مدل عمومی رگرسیون خطی ۱۱۰
تحلیل واریانس برای بررسی برازش مدل رگرسیون خطی چندگانه ۱۱۳
استنباط در مورد پارامترهای مدل رگرسیون چندگانه ۱۱۵
تمرینات ۱۱۶
فصل چهارم: آزمون‌های ناپارمتری ۱۱۹
مقدمه ۱۲۰
آزمون‌های ناپارامتری یک نمونه‌ای ۱۲۰
آزمون گردش ۱۲۰
آزمون دوجمله‌ای ۱۲۴
آزمون مجذور کای برای نیکویی برازش ۱۲۶

آزمون نیکویی برازش کولموگروف- اسمیرنوف ۱۲۸
جداول توافقی 2×2 ..... ۱۳۱
جداول توافقی $r\times c$ (آزمون همگنی) ۱۳٤
آزمون استقلال دو متغیر کیفی ۱۳۷
تمرینات ۱۳۹
آزمون‌های همبستگی ۱٤۱
ضریب همبستگی رتبه‌ای اسپیرمن ۱٤۱
تعدیل اثر داده‌های تکراری ۱٤۳
آزمون ضریب همبستگی اسپیرمن ۱٤٤
ضریب همبستگی گاما ۱٤٥
آزمون ضریب همبستگی رتبه‌ای گاما ۱٤۷
ضریب همبستگی فی (ϕ) ۱٤۷
آزمون ضریب همبستگی فی (ϕ) ۱٤۸
تمرینات ۱٤۹
آزمون‌های ناپارامتری در مورد مقایسه دو جامعه ۱٥۱
آزمون‌های مربوط به دو جامعه مستقل ۱٥۱
آزمون دو نمونه‌ای کولموگروف- اسمیرنوف ۱٥۱
آزمون من - ویتنی ۱٥٤
آزمون مک نمار برای مقایسه نسبتها در دو جامعه وابسته ۱٥۷
آزمون علامت برای دو نمونه وابسته ۱٥۹
آزمون رتبه علامتدار ویلکاکسون ۱٦۱
آزمون کروسکال- والیس برای مقایسه بیش از دو جامعه مستقل ۱٦۳
آزمون فریدمن برای k نمونه وابسته ۱٦٦
آزمون تعقیبی فریدمن ۱٦۷
آزمون کوکران ۱٦۹
تمرینات ۱۷۱

قسمت دوم: معرفی نرم‌افزار R

مقدمه ۱۷۶

فصل پنجم: مقدمات R ۱۷۷

مقدمه ۱۷۸

محیط R ۱۷۸

شروع کار با R ۱۷۹

بسته‌ها ۱۸۰

دریافت کمک از منوی HELP ۱۸۱

توابع و شناسه ها ۱۸۳

بردار ۱۸۴

بردارهای محاسباتی ۱۸۵

ایجاد توالی ۱۸۶

ماتریس ۱۸۸

جدول ۱۹۱

جدول یک طرفه ۱۹۱

جدول دو طرفه ۱۹۲

چارچوب داده ۱۹۳

انتخاب زیر مجموعه‌ای از داده‌ها ۱۹۴

نحوه ورود داده به محیط R ۱۹۶

فصل ششم: آمار توصیفی ۱۹۷

نمودارها ۲۰۰

هیستوگرام ۲۰۲

نمودار جعبه‌ای ۲۰۷

نمودار پراکنش ۲۱۰

بررسی نرمال بودن داده‌ها ۲۱۱

نمودار میله‌ای ۲۱۳

نمودار دایره‌ای ۲۱۶

نمودار ساقه و برگ ۲۱۷
نمایش داده‌های چند متغیره ۲۱۸
نمایش توأم نمودارها ۲۱۹
فصل هفتم: احتمال، نمونه‌گیری، برآورد ۲۲۱
نمونه برداری تصادفی ۲۲۲
محاسبه احتمال ۲۲۳
توزیع های موجود در R ۲۲٤
ایجاد فاصله اطمینان برای میانگین ۲۲۷
فصل هشتم: آمار تحلیلی ۲۲۹
آزمون‌های یک و دو نمونه‌ای ۲۳۰
آزمون T یک نمونه‌ای ۲۳۰
آزمون ویلکاکسون برای میانگین یک جامعه (یک نمونه‌ای) ۲۳۳
آزمون T دونمونه‌ای ۲۳٤
مقایسه واریانس‌ها ۲۳٦
آزمون ویلکاکسون دو نمونه‌ای ۲۳۷
آزمون T زوجی ۲۳۷
آزمون زوجی ویلکاکسون ۲۳۸
آنالیز واریانس و کروسکالوالیس ۲۳۹
آنالیز واریانس یک عاملی ۲۳۹
آزمون بارتلت ۲٤۲
آزمون کروسکال والیس ۲٤۳
آنالیز واریانس دوعاملی ۲٤۳
نمودار اثر متقابل ۲٤٥
آزمون فریدمن ۲٤٦
رگرسیون ۲٤٦
همبستگی ۲٤۸
ضریب همبستگی پیرسون ۲٤۸

ضریب همبستگی اسپیرمن..... ۲۵۰
ضریب همبستگی کندال..... ۲۵۱
فهرست منابع..... ۲۵۳
ضمائم..... ۲۵۹

فصل اول

برآورد

مقدمه

یکی از اهداف بررسی های آماری یافتن تخمینی برای پارامترهای مجهول جامعه است. این کار بایستی با دقت کافی صورت گیرد تا بتوان به تخمین ارائه شده تا حد قابل توجهی اطمینان داشت. پارامتر شاخص جامعه است که مقدار آن ثابت می‌باشد و آماره شاخص مربوط به نمونه است که از یک نمونه به نمونه‌ی دیگر تغییر می‌کند. مثلاً میانگین یک جامعه پارامتر و میانگین نمونه‌ای از آن جامعه آماره می‌باشد. معمولاً برای تخمین یک پارامتر ابتدا نمونه‌ای تصادفی از جامعه مورد بررسی انتخاب نموده و سپس با استفاده از این نمونه، آماره‌ای به دست می‌آوریم که به عنوان تخمینی برای پارامتر مجهول استفاده می‌شود. این تخمین بایستی دارای ویژگیهایی باشد که در این فصل به این ویژگیها می‌پردازیم.

تعریف برآورد

تخمین پارامتر مجهول جامعه با استفاده از آماره نمونه را برآورد می‌نامیم.

به عنوان مثال اگر میانگین جامعه‌ای مجهول باشد برای تخمین آن یک نمونه تصادفی از جامعه انتخاب و میانگین این نمونه را به عنوان تخمینی برای میانگین جامعه استفاده می‌کنیم. در این بخش پارامتر را با θ و آماره را با $\hat{\theta}$ نمایش می دهیم.

خواص برآورد کننده‌ها

نااریبی: گوئیم $\hat{\theta}$ یک برآورد کننده نااریب برای θ است هرگاه:

$$E(\hat{\theta}) = \theta$$

مقدار $B(\hat{\theta}) = E(\hat{\theta}) - \theta$ را مقدار اریبی برآورد کننده θ می‌نامند.