

روش‌های آماری برای متخصصان کتابداری و اطلاع‌رسانی

رویکرد کاربردی و ساده برای درک،
استفاده و تفسیر علم آمار

نوشته

لیون واگان

مترجمان

محمد رضا قانع - کیوان کوشا

ویراستار

دکتر حمیدرضا حسن آبادی



Vaughan, Liwen

م

واگان لیون، ۱۹۵۶ -

روش‌های آماری برای متخصصان کتابداری و اطلاع‌رسانی: رویکرد کاربردی و ساده
برای درک استفاده و تفسیر علم آمار / نوشته لیون وان؛ مترجمان محمدرضا قانع، کیوان
کوشا؛ ویراستار دکتر حمیدرضا حسن‌آبادی. - تهران: چاپار، ۱۳۸۴.
۲۷۲ ص.

ISBN : 964-7790-03-1

فهرست‌نویسی براساس اطلاعات فیپا.

عنوان اصلی: Statistical methods for the information professional 2001.

۱. آمار ریاضی. ۲. کتابداری - تحقیق - روش‌های آماری. ۳. اطلاع‌رسانی - تحقیق -
روش‌های آماری. ۴. آمار بازرگانی. الف. قانع، محمدرضا، ۱۳۳۸ - ، مترجم. ب. کوشا،
کیوان، ۱۳۳۷ - ، مترجم. ج. عنوان.

۵۱۹/۵

QA۲۷۶/۲۴۹

۱۳۸۴

۸۴-۵۰۵۰م

کتابخانه ملی ایران



تهران: ۸۸۹۹۶۸۰ - ۹۱۲۱۹۶۰۲۱۴

روش‌های آماری برای متخصصان کتابداری و اطلاع‌رسانی

رویکرد کاربردی و ساده برای درک استفاده و تفسیر علم آمار

نوشته لیون واگان

مترجمان: محمدرضا قانع - کیوان کوشا

ویراستار: دکتر حمیدرضا حسن‌آبادی

چاپ اول ۱۳۸۴

شمارگان: ۲۰۰۰ نسخه

۳۰۰۰۰ ریال

حق چاپ محفوظ است

ISBN : 964-7790-03-1

شابک: ۹۶۴-۷۷۹۰-۰۳-۱

فهرست مطالب

پیشگفتار

- کتابی متمایز در حوزه آمار ۹
- چه کسانی باید این کتاب را مطالعه کنند؟ ۱۱
- چگونه از این کتاب استفاده شود؟ ۱۱
- پیشگفتار مترجمان ۱۲

فصل اول

- شروع کار - تشخیص انواع داده‌ها ۱۵
- ۱-۱ داده‌های اسمی ۱۶
- ۲-۱ داده‌های رتبه‌ای ۱۷
- ۳-۱ داده‌های فاصله‌ای ۱۸
- ۴-۱ داده‌های نسبی ۲۰
- ۵-۱ تبدیل داده‌ها ۲۱

فصل دوم

- پرهیز از محاسبات دستی و بکاربردن فرمول‌ها - استفاده از نرم‌افزارهای آماری ۲۷
- ۱-۲ انواع نرم‌افزارهای آماری ۲۸

- ۲-۲ چه نرم افزاری را انتخاب کنیم؟ ۲۹
- ۳-۲ چگونه داده‌ها را در فایل‌های کامپیوتری سازماندهی کنیم؟ ۳۱
- ۴-۲ چگونه باید با داده‌های مفقود شده مواجه شد؟ ۳۲

فصل سوم

- گام نخست - استفاده از نمودارها به منظور مشاهده خصوصیات داده‌ها ۳۷
- ۱-۳ انواع نمودارها ۳۸
- ۲-۳ یک نمودار میله‌ای مخصوص - (هیستوگرام) ۴۴

فصل چهارم

- خلاصه کردن داده‌های نامنظم به صورت اعداد منظم - آمار توصیفی ۵۱
- ۱-۴ اندازه‌های گرایش مرکزی ۵۲
- ۱-۱-۴ میانگین - معدل حسابی ۵۲
- ۲-۱-۴ میانه - نقطه وسط ۵۵
- ۳-۱-۴ نما - راس هیستوگرام ۵۶
- ۴-۱-۴ چه موقع از کدام اندازه گرایش مرکزی استفاده کنیم؟
- خلاصه ۵۹
- ۲-۴ اندازه‌های پراکندگی ۶۱
- ۱-۲-۴ دامنه تغییرات ۶۲
- ۲-۲-۴ دامنه میان چارکی ۶۴
- ۳-۲-۴ انحراف معیار ۶۴
- ۴-۲-۴ واریانس ۶۸
- ۵-۲-۴ چه موقع از کدام اندازه پراکندگی استفاده کنیم؟
- خلاصه ۶۸
- ۳-۴ ارتباط دادن اندازه‌های آمار توصیفی به یکدیگر - مثال‌ها ۶۹

فصل پنجم

- معنی داری آماری چیست؟ - مفاهیم بنیادی آمار استنباطی ۷۷
- ۱-۵ آمار توصیفی در برابر آمار استنباطی ۷۸
- ۲-۵ جامعه در برابر نمونه ۷۸
- ۳-۵ پارامتر در برابر آماره ۷۹
- ۴-۵ احتمالات و توزیع فراوانی ۸۱
- ۵-۵ توزیع نرمال ۸۴
- ۶-۵ نمره Z ۸۵
- ۷-۵ توزیع نرمال معیار ۸۸
- ۸-۵ فاصله اطمینان ۹۰
- ۹-۵ فرض آزمایی - معنی داری یا عدم معنی داری ۹۶
- ۱۰-۵ خطاهای آزمون آماری - خطای نوع اول و خطای نوع دوم ۱۰۳

فصل ششم

- چگونه داده‌ها را جمع آوری کنیم - روش‌های نمونه‌گیری ۱۰۷
- ۱-۶ نمونه‌گیری تصادفی ساده ۱۰۸
- ۲-۶ نمونه‌گیری سیستماتیک ۱۱۰
- ۳-۶ نمونه‌گیری طبقه‌ای ۱۱۱
- ۴-۶ نمونه‌گیری اریب ۱۱۳

فصل هفتم

- بررسی روابط داده‌های اسمی و رتبه‌ای آزمون مجذور کای ۱۱۹
- ۱-۷ منطق آزمون مجذور کای ۱۲۰
- ۲-۷ محاسبه فراوانی‌های مورد انتظار ۱۲۳
- ۳-۷ نمره مجذور کای ۱۲۵

- ۴-۷ جدول مجذور کای ۱۲۷
- ۵-۷ بررسی الگوی رابطه ۱۳۱
- ۶-۷ مثالی در مورد استفاده از نرم افزار برای انجام یک آزمون مجذور کای ۱۳۲
- ۷-۷ شرایط استفاده از آزمون مجذور کای ۱۳۷

فصل هشتم

- بررسی روابط برای داده‌های فاصله‌ای و نسبی - همبستگی و رگرسیون ۱۴۳
- ۱-۸ انواع همبستگی ۱۴۴
- ۲-۸ استفاده از نمودار پراکندگی برای نمایش الگوی ارتباط میان داده‌ها ۱۴۵
- ۳-۸ اندازه‌گیری شدت ارتباط - ضریب همبستگی پیرسون ۱۴۸
- ۴-۸ آزمون معنی‌داری ضریب همبستگی پیرسون ۱۵۱
- ۵-۸ همبستگی و علّیت ۱۵۴
- ۶-۸ معادله و خط رگرسیون ۱۵۶
- ۷-۸ پیش‌بینی ۱۶۳
- ۸-۸ شرایط لازم برای انجام تحلیل همبستگی و رگرسیون ۱۶۶

فصل نهم

- آیا دو نمونه تفاوت معنی‌داری دارند؟ آزمون T ۱۶۹
- ۱-۹ آزمون T مستقل در برابر آزمون T جفتی ۱۷۰
- ۲-۹ منطق آزمون T ۱۷۴
- ۳-۹ شیوه انجام آزمون T ۱۷۷
- ۴-۹ مثال‌هایی از آزمون T با استفاده از نرم افزار ۱۷۹
- ۵-۹ شرایط لازم برای استفاده از آزمون T ۱۸۴

فصل دهم

- آیا بین سه نمونه یا بیشتر اختلاف معنی داری وجود دارد؟ تحلیل واریانس ۱۸۷
- ۱-۱۰ منطق تحلیل واریانس (ANOVA) ۱۸۹
- ۲-۱۰ شیوه انجام تحلیل واریانس ۱۹۳
- ۳-۱۰ مثالی از تحلیل واریانس با استفاده از نرم افزار ۱۹۵
- ۴-۱۰ بررسی الگوی تفاوت ۱۹۸
- ۵-۱۰ شرایط لازم برای استفاده از تحلیل واریانس ۲۰۳

فصل یازدهم

- استفاده از آزمون های آماری غیرپارامتریک ۲۰۷
- ۱-۱۱ ضریب همبستگی اسپیرمن ۲۰۹
- ۲-۱۱ آزمون مان - ویتنی ۲۱۴
- ۳-۱۱ آزمون رتبه های علامت دار ویلکاکسون ۲۱۷
- ۴-۱۱ آزمون کروسکال - والیس ۲۲۱
- ۵-۱۱ مزایا و محدودیت های آزمون های غیرپارامتریک ۲۲۶
- ۱-۵-۱۱ مزایای آزمون های غیرپارامتریک ۲۲۶
- ۲-۵-۱۱ محدودیت های آزمون های غیرپارامتریک ۲۲۷
- ۳-۵-۱۱ چه زمانی از آزمون های غیرپارامتریک استفاده کنیم؟ ۲۲۹

فصل دوازدهم

- چه زمانی از چه آزمونی استفاده کنیم؟ - نقشه جهت یابی ۲۳۱

فصل سیزدهم

- کسب مهارت بیشتر - تجربه ای از چند روش آماری پیشرفته ۲۳۹
- ۱-۱۳ تحلیل واریانس دوعاملی ۲۴۰

- ۲-۱۳ رگرسیون چندگانه ۲۴۹
- ۱-۲-۱۳ چرا به رگرسیون چندگانه نیاز داریم؟ ۲۴۹
- ۲-۲-۱۳ معادله رگرسیون چندگانه ۲۵۱
- ۳-۲-۱۳ ضرایب رگرسیون ۲۵۳
- ۴-۲-۱۳ ضریب همبستگی چندگانه و ضریب تعیین چندگانه ۲۵۴
- ۵-۲-۱۳ ضریب همبستگی جزئی ۲۵۵
- ۳-۱۳ روابط ساختاری خطی ۲۵۶

ضمیمه‌ها

- ضمیمه ۱: توزیع نرمال معیار ۲۶۵
- ضمیمه ۲: جدول اعداد تصادفی ۲۶۶
- ضمیمه ۳: مقادیر بحرانی مجذور کای ۲۶۷
- ضمیمه ۴: مقادیر بحرانی ضریب همبستگی پیرسون ۲۶۸
- ضمیمه ۵: مقادیر بحرانی آزمون T ۲۶۹
- ضمیمه ۶: مقادیر بحرانی آزمون F برای تحلیل واریانس ($\alpha = 0.05$) ۲۷۰
- ضمیمه ۷: مقادیر بحرانی برای آزمون توکی ($\alpha = 0.05$) ۲۷۱

کتابشناسی ۲۷۳

معرفی مولف ۲۷۵

پیشگفتار

کتابی متمایز در حوزه آمار

نکته قابل توجه این که آمار در تحلیل داده‌ها یکی از مفیدترین و قدرتمندترین ابزارها است و همچنین یکی از نگران‌کننده‌ترین و دوست‌نداشتنی‌ترین موضوعات تحقیق است. ریشه این نگرانی در این است که کتاب‌های آمار معمولاً آکنده از معادلات، زبان نامفهوم ریاضی، صفحاتی از مشتقات و حساب هستند. آنچه این کتاب را از کتاب‌های دیگر متمایز می‌کند تأکید آن بر درک منطق پایه آمار است تا پیچیدگی‌های ریاضی. توجه در این کتاب بر مفهوم آمار است، چه زمانی آنها را بکار ببریم، چگونه آنها را بکار ببریم، و چگونه نتایج را تفسیر کنیم. این کتاب نشان دهنده روش موفقیت‌آمیزی است که طی ده سال تدریس درس آمار کسب نموده‌ام. سه قسمت برای این روش وجود دارد:

اول، برای توضیح مفاهیم و آزمون‌های آماری از استدلال منطقی استفاده کرده‌ام تا استنباط ریاضی. در حالی که علم آمار شاخه‌ای از علم ریاضی است، اما منطق زیر بنایی بیشتر آزمون‌های آماری به قدری ساده است که برای درک آن نیازی به سابقه ریاضی پیشرفته نمی‌باشد. معادلات و توضیحات ریاضی که معمولاً در کتابهای آمار هستند اغلب سادگی این منطق زیربنایی را می‌پوشانند. از این رو، بسیاری از این

معادلات را در این کتاب مشاهده نمی‌کنید. آن تعداد کمی هم که ملاحظه می‌کنید با دقت توضیح داده شده‌اند و بدون اینکه به درک شما از اصول اولیه مورد بحث لطمه‌ای وارد کند می‌تواند نادیده گرفته شود.

دوم، به منظور نشان دادن فرایند هر آزمون آماری مثالی از یک تحقیق در حوزه علم اطلاع‌رسانی ارائه داده‌ام. با چگونگی منظم کردن فرضیه‌ها شروع کرده‌ام، سپس استفاده از نرم افزار کامپیوتری را جهت تحلیل داده‌ها، تفسیر برون‌داد، و دست یافتن به نتایج را گزارش نموده‌ام. خوانندگان بیشتر تصویر کاملی از چگونگی بکارگیری روش‌های آماری بدست می‌آورند تا اینکه محو جزئیات فنی شوند. تاکید بر روی کاربردهای واقعی روش‌های آماری خواننده را درک این نکته که چگونه این روش‌ها نسبت به قسمت پنهانی ریاضیات که نامربوط هستند، به عنوان ابزاری قدرتمند و مربوط محقق را کمک می‌کند. سرانجام، برای محاسبات و کارهای سخت ریاضی از نرم افزار کامپیوتری استفاده کرده‌ام. بسیاری از کتابهای آمار وقت زیادی را برای ساختن فرمول‌ها و انجام محاسبات صرف می‌کنند. واقعیت این است که امروزه هیچ کس این محاسبات را به صورت دستی انجام نمی‌دهد. بسته‌های نرم‌افزاری کامپیوتری متنوعی وجود دارد که می‌تواند برای انجام محاسبات سخت ریاضی مورد استفاده قرار گیرند، بنابراین می‌توانیم روی استنباط نتایج تاکید کنیم. به منظور شناخت این مطلب، یک فصل را به مسائلی که مخصوص استفاده از نرم افزار کامپیوتری برای انجام تحلیل‌های آماری است اختصاص داده‌ام.

از این روش برای تدریس آمار به دانشجویان رشته‌های مختلف در مقاطع لیسانس تا دکتری بهره‌جسته‌ام. اغلب دانشجویان در کلاس‌های درس من از پیشینه ریاضی بسیار اندکی برخوردار بوده‌اند و بیشتر آنها با ترس و نگرانی این درس را انتخاب کرده‌اند. با این حال، به نحو خوشایندی از اینکه آمار چقدر آسان است شگفت زده بودند. چقدر جالب است موضوعی را که زمانی برای آنها اسرار آمیز و نامانوس بوده

است، می‌فهمند. برآستی، اگر آمار را مقوله‌ای پیچیده و خارج از فهم مردمی که ریاضیدان نیستند، بدانیم چیزی است شبیه یک افسانه. مطمئن هستیم زمانی که این کتاب را مطالعه کنید با من هم عقیده خواهید شد.

چه کسانی باید این کتاب را مطالعه کنند؟

متخصصان کتابداری و اطلاع‌رسانی، هم اساتید و هم متخصصان، این کتاب را برای درک مفاهیم آماری و روش‌هایی که در علم اطلاع‌رسانی بسیار استفاده می‌شود و به وفور در نوشته‌ها ظاهر می‌گردد، مفید می‌یابند. شما به عنوان یک متخصص اطلاع‌رسانی، حداقل یک استفاده‌کننده از نتایج آماری تولید شده توسط دیگران هستید اگر چه خود تولیدکننده آمار نمی‌باشید. هدف کتاب حاضر این است که منبعی را در اختیار شما بگذارد که به شما هم در فهم و هم در بکارگیری روش‌های آماری کمک نماید. اگر، شما یک متخصص اطلاع‌رسانی نیستید، باز هم با مطالعه این کتاب می‌توانید از آن بهره‌مند شوید. با این همه، اصول و نظریه‌های آمار در همه رشته‌های علمی یکسان است.

چگونه از این کتاب استفاده شود؟

این کتاب روش‌های آماری پایه را که بطور معمول در تحقیقات حوزه کتابداری و اطلاع‌رسانی مورد استفاده است پوشش می‌دهد. این روش‌ها شامل آمار استنباطی، آزمون t ، آزمون مجذور کای، تحلیل واریانس، همبستگی، رگرسیون و آزمون‌های پایه آماری غیرپارامتریک است. این کتاب همچنین بعضی از آزمون‌های آماری متوسط و پیشرفته مانند رگرسیون چندگانه و LISREL را معرفی می‌کند.

چنانچه از قبل با تحلیل آماری آشنائی دارید و از این کتاب به منظور یادآوری موضوعی خاص استفاده می‌کنید، می‌توانید مستقیماً به بخش مورد علاقه مراجعه نمایید. بخصوص، نمودار موجود در فصل ۱۲ را می‌توانید به منظور انتخاب آزمون مورد نظر و مرور آن برای یک موقعیت مشخص ملاحظه نمایید و سپس می‌توانید به فصل مربوطه مراجعه کنید.

اگر برای اولین مرتبه است که به یادگیری آمار مبادرت می‌کنید یا اگر آن را قبلاً یاد گرفته‌اید ولی نیاز به بازآموزی مفاهیم اصلی دارید، توصیه می‌کنم که قبل از مطالعه فصل‌های دیگر ابتدا شش فصل اول را به ترتیب مطالعه نمایید. فصل اول به چگونگی تشخیص انواع مختلف داده می‌پردازد، اولین موضوعی است که در آغاز تحلیل داده‌ها با آن مواجه می‌شوید. فصل دوم مسائل کلی در رابطه با استفاده از نرم افزار کامپیوتری در تحلیل آماری را ارائه می‌دهد، که شامل چگونگی سازماندهی داده‌ها در فایل‌های کامپیوتری و چگونگی برخورد با داده‌های مفقود شده است. فصل‌های سوم و چهارم به ترتیب نمودارها و آمار توصیفی را بحث می‌کند. فصل پنجم مفاهیم اولیه آمار استنباطی را معرفی می‌نماید. چون این فصل اساس آزمون‌های آماری بحث شده در فصل‌های بعدی را مطرح می‌کند، بنابراین قبل از مراجعه به فصل‌های بعدی باید مورد مطالعه قرار گیرد. فصل ششم شامل انواع روش‌های نمونه‌گیری است.

اگرچه ماهیت موضوعی شش فصل اول به گونه‌ای است که باید به ترتیب مطالعه شوند، اما فصل‌های هفتم تا یازدهم به این روش مطالعه نمی‌شوند، زیرا هر فصل آزمون آماری متفاوتی را معرفی می‌کند. بنابراین می‌توانید مستقیماً به فصل مورد علاقه مراجعه نمایید. فصل دوازدهم آزمون‌های آمار استنباطی فصل‌های هفتم تا یازدهم را خلاصه می‌کند. نموداری به منظور راهنمایی شما به آزمون آماری مورد نظر جهت استفاده در شرایط خاص طراحی شده است. فصل سیزدهم مقدمه‌ای است بر روش‌های آماری ابتدائی تا پیشرفته. به خاطر اینکه بیشتر مفاهیم ارائه شده در آن در فصل‌های قبل معرفی شده، توصیه می‌شود که این فصل در آخر مورد مطالعه قرار گیرد.

پیشگفتار مترجمان

استفاده از الگوها و روش‌های آماری برای تحلیل داده‌ها یک فعالیت معمول در تمام رشته‌های علمی است. مفاهیم و روش‌های آماری برای درک جهان اطراف ما حیاتی می‌باشد. علم آمار به ما می‌آموزد که چگونه یک قضاوت عاقلانه و تعمیمی آگاهانه در برابر عدم قطعیت اتخاذ نمائیم. نقش حیاتی آمار در پژوهش‌های علمی دست‌اندرکاران این حوزه را مجبور به فراگیری این علم نموده است. در سال‌های اخیر آموزش آمار دستخوش تغییراتی شده است. تجهیزات کامپیوتری و نرم‌افزارهای آماری محاسبات پرزحمت آماری را حذف کرده است و امروزه بیشتر توجهات به مسئله تحلیل و تفسیر نتایج معطوف گردیده است.

محققان حوزه کتابداری و اطلاع‌رسانی همانند دیگر حوزه‌های علوم در فرایند انجام پژوهش مبادرت به جمع‌آوری داده‌های عددی درخصوص یک موضوع می‌کنند. کسب دانش روش‌های تحلیل کتبی، متخصصان را در تفسیر صحیح این داده‌ها یاری خواهد کرد. بوشاوارتر اظهار می‌دارند که استفاده از آمار کتابداران را در بهره‌گیری از حداکثر اطلاعات نتایج تحقیقات کمک می‌کند. لذا، کتاب حاضر با همین انگیزه و به لحاظ معرفی علم آمار از رویکردی متفاوت یعنی درک و فهم مفاهیم آماری (از نوع توصیفی و استنباطی) به جای توجه به فرمول‌های پیچیده ریاضی ترجمه شده است که امید است مورد قبول خوانندگان بویژه متخصصان حوزه کتابداری و اطلاع‌رسانی واقع شود.

اکنون که ترجمه اثر حاضر به زیور طبع آراسته شده است مراتب سپاس خود را از اساتید بزرگوار دکتر عباس بازرگان و دکتر نوروزعلی کرم‌دوست برای پیوند دوباره دانشجویان خود با علم آمار که خود انگیزه دیگر ترجمه کتاب حاضر شد، ابرار می‌داریم. همچنین از همکار محترم دکتر حمید حسن‌آبادی به خاطر دقت و حوصله ایشان در ویرایش علمی اثر حاضر سپاسگزاریم. در آخر از آقای کریم فلاح برای مهارت و دقت در تایپ رایانه‌ای این کتاب تشکر می‌نمائیم.

شایان ذکر است، کتاب حاضر دارای اجازه چاپ از مولف کتاب و انجمن اطلاع‌رسانی و فناوری اطلاعات آمریکا (asis & t) می‌باشد.

محمدرضا قانع - کیوان کوشا

بهار ۱۳۸۴