



دانشگاه سям نور

آمار در کتابداری و اطلاع رسانی

(کارشناسی ارشد علم اطلاعات و دانش شناسی)

دکتر پرویز نصیری دکتر ثریا ضیائی دکتر باقر مقدس زاده

امروزه کتاب‌خوانی و علم‌آموزی، نه تنها یک وظیفه‌ی ملی، که یک واجب دینی است.

مقام معظم رهبری

در عصر حاضر یکی از شاخصه‌های ارزیابی رشد، توسعه و پیشرفت فرهنگی هر کشوری میزان تولید کتاب، مطالعه و کتاب‌خوانی مردم آن مرز و بوم است. ایران اسلامی نیز از دیرباز تاکنون با داشتن تمدنی چندهزارساله و مراکز متعدد علمی، فرهنگی، کتابخانه‌های معتبر، علما و دانشمندان بزرگ با آثار ارزشمند تاریخی، سرآمد دولت‌ها و ملت‌های دیگر بوده و در عرصه‌ی فرهنگ و تمدن جهانی به‌سان خورشیدی تابناک همچنان می‌درخشد و با فرزندان نیک‌نهاد خویش هنرنمایی می‌کند. چه کسی است که در دنیا با دانشمندان فرزانه و نام‌آور ایرانی همچون ابوعلی سینا، ابوریحان بیرونی، فارابی، خوارزمی و ... همچنین شاعران برجسته‌ای نظیر فردوسی، سعدی، مولوی، حافظ و ... آشنا نباشد و در مقابل عظمت آنها سر تعظیم فرود نیاورد. تمامی این افتخارات ارزشمند، برگرفته از میزان عشق و علاقه فراوان ملت ما به فراگیری علم و دانش از طریق خواندن و مطالعه منابع و کتاب‌های گوناگون است. به شکرانه‌ی الهی، تاریخ و گذشته ما، همیشه درخشان و پربار است. ولی اکنون در این زمینه در چه جایگاهی قرار داریم؟ آمار و ارقام ارائه‌شده از سوی مجامع و سازمان‌های فرهنگی در مورد سرانه‌ی مطالعه‌ی هر ایرانی، برایمان چندان امیدوارکننده نمی‌باشد و رهبر معظم انقلاب اسلامی نیز از این وضعیت بارها اظهار گله و ناخشنودی نموده‌اند.

کتاب، دروازه‌ای به سوی گستره‌ی دانش و معرفت است و کتاب خوب، یکی از بهترین ابزارهای کمال بشری است. همه‌ی دستاوردهای بشر در سراسر عمر جهان تا آنجا که قابل کتابت بوده است، در میان دست‌نوشته‌هایی است که انسان‌ها پدید آورده و می‌آورند. در این مجموعه‌ی بی‌نظیر، تعالیم الهی، درس‌های پیامبران به بشر، و همچنین علوم مختلفی است که سعادت بشر بدون آگاهی از آنها امکان‌پذیر نیست. کسی که با دنیای زیبا و زندگی‌بخش کتاب ارتباط ندارد بی‌شک از مهم‌ترین دستاوردهای انسانی و نیز از بیشترین معارف الهی و بشری محروم است. با این دیدگاه، به‌روشنی می‌توان ارزش و مفهوم رمزی عمیق در این حقیقت تاریخی را دریافت که اولین خطاب خداوند متعال به پیامبر گرامی اسلام (ص) این است که «بخوان!» و در اولین

سوره‌ای که بر آن فرستاده‌ی عظیم‌الشان خداوند، فرود آمده، نام «قلم» به تجلیل یاد شده‌است: «اقْرَأْ وَ رَبُّكَ الْأَكْرَمُ. الَّذِي عَلَّمَ بِالْقَلَمِ» در اهمیت عنصر کتاب برای تکامل جامعه‌ی انسانی، همین بس که تمامی ادیان آسمانی و رجال بزرگ تاریخ بشری، از طریق کتاب جاودانه مانده‌اند.

دانشگاه پیام‌نور با گستره‌ی جغرافیایی ایران‌شمول خود با هدف آموزش برای همه، همه‌جا و همه‌وقت، به‌عنوان دانشگاهی کتاب‌محور در نظام آموزش عالی کشورمان، افتخار دارد جایگاه اندیشه‌سازی و خردورزی بخش عظیمی از جوانان جویای علم این مرز و بوم باشد. تلاش فراوانی در ایام طولانی فعالیت این دانشگاه انجام پذیرفته تا با بهره‌گیری از تجربه‌های گرانقدر استادان و صاحب‌نظران برجسته کشورمان، کتاب‌ها و منابع آموزشی درسی شاخص و خودآموز تولید شود. در آینده هم، این مهم با هدف ارتقای سطح علمی، روزآمدی و توجه بیشتر به نیازهای مخاطبان دانشگاه پیام‌نور با جدیت ادامه خواهد داشت. به‌طور قطع استفاده از نظرات استادان، صاحب‌نظران و دانشجویان محترم، ما را در انجام این وظیفه‌ی مهم و خطیر یاری‌رسان خواهد بود. پیشاپیش از تمامی عزیزانی که با نقد، تصحیح و پیشنهادهای خود ما را در انجام این وظیفه‌ی خطیر یاری می‌رسانند، سپاسگزاری می‌نماییم. لازم است از تمامی اندیشمندانی که تاکنون دانشگاه پیام‌نور را منزلگاه اندیشه‌سازی خود دانسته و ما را در تولید کتاب و محتوای آموزشی درسی یاری نموده‌اند، صمیمانه قدردانی گردد. موفقیت و بهروزی تمامی دانشجویان و دانش‌پژوهان عزیز آرزوی همیشگی ما است.

دانشگاه پیام‌نور

فهرست مطالب

پیشگفتار.....	نه
فصل اول. کلیات و مفاهیم اولیه.....	۱
هدف کلی.....	۱
هدف‌های یادگیری.....	۱
مقدمه.....	۱
۱-۱ حوزه فعالیت‌های کتاب‌سنجی و کاربرد آمار.....	۳
۲-۱ مهم‌ترین مشخصه‌های کتاب‌سنجی.....	۴
۳-۱ کتاب‌سنجی نوین در ایران.....	۴
۴-۱ مطالعات بین‌المللی در حوزه‌های سنجشی.....	۷
۵-۱ مفاهیم اولیه آمار توصیفی.....	۹
۱-۵-۱ جامعه.....	۹
۲-۵-۱ نمونه.....	۱۱
۳-۵-۱ روش‌های نمونه‌گیری.....	۱۱
۴-۵-۱ متغیر.....	۱۲
۵-۵-۱ مقیاس‌های اندازه‌گیری.....	۱۴
خلاصه فصل اول.....	۱۶
خودآزمایی تشریحی فصل اول.....	۱۶
فصل دوم. طبقه‌بندی داده‌ها و محاسبه شاخص‌های آماری.....	۱۹
هدف کلی.....	۱۹
هدف‌های یادگیری.....	۱۹
مقدمه.....	۱۹
۱-۲ جدول‌های آماری.....	۲۰

۲۱	۱-۱-۲ جدول توزیع فراوانی برای داده‌های گسسته
۲۲	۲-۱-۲ جدول توزیع فراوانی برای داده‌های پیوسته
۲۷	۲-۲ نمودارهای آماری
۲۷	۱-۲-۲ نمودارهای آماری برای داده‌های گسسته
۳۰	۲-۲-۲ نمودارهای آماری برای داده‌های پیوسته
۳۴	۳-۲ شاخص‌های مرکزی و پراکندگی
۳۴	۱-۳-۲ شاخص‌های مرکزی
۴۷	۲-۳-۲ شاخص‌های پراکندگی
۵۷	خلاصه فصل دوم
۵۷	خودآزمایی تشریحی فصل دوم
۶۱	فصل سوم. متغیرهای تصادفی و توزیع‌های مهم آماری
۶۱	هدف کلی
۶۱	هدف‌های یادگیری
۶۱	مقدمه
۶۲	۱-۳ متغیر تصادفی
۶۳	۲-۳ توزیع احتمالات گسسته
۶۴	۳-۳ توزیع احتمالات پیوسته
۶۶	۴-۳ تابع توزیع
۶۸	۱-۴-۳ توزیع یکنواخت گسسته
۶۸	۲-۴-۳ توزیع برنولی
۶۹	۳-۴-۳ توزیع دو جمله‌ای
۷۰	۴-۴-۳ توزیع پواسن
۷۲	۵-۳ توزیع یکنواخت پیوسته
۷۲	۶-۳ توزیع نرمال
۷۵	۷-۳ نحوه استفاده از جدول Z
۷۸	۸-۳ توزیع میانگین نمونه‌ها و برآورد نقطه‌ای میانگین جامعه
۸۰	۹-۳ توزیع استودنت ^t
۸۲	۱۰-۳ امید ریاضی
۸۳	۱۱-۳ امید ریاضی تابعی از X
۸۴	۱۲-۳ ویژگی‌های امید ریاضی
۸۵	۱۳-۳ واریانس جامعه
۸۶	۱۴-۳ ویژگی‌های واریانس
۸۷	خلاصه فصل سوم
۸۷	خودآزمایی تشریحی فصل سوم

فصل چهارم. آزمون فرضیه آماری	۹۱
هدف کلی	۹۱
هدف‌های یادگیری	۹۱
مقدمه	۹۱
۱-۴ مفاهیم پایه در آزمون فرض	۹۲
۲-۴ آزمون فرض درباره میانگین جامعه نرمال	۹۴
۳-۴ آزمون فرض درباره نسبت جامعه	۹۹
خلاصه فصل چهارم	۱۰۲
خودآزمایی تشریحی فصل چهارم	۱۰۳
فصل پنجم. هم‌بستگی و انواع آن	۱۰۵
هدف کلی	۱۰۵
هدف‌های یادگیری	۱۰۵
مقدمه	۱۰۵
۱-۵ انواع هم‌بستگی	۱۰۶
۱-۵-۱ نمودار پراکنش	۱۰۸
۲-۵ ضریب هم‌بستگی	۱۰۹
۳-۵ ضریب هم‌بستگی پیرسون	۱۰۹
۴-۵ ضریب تعیین	۱۱۳
۵-۵ هم‌بستگی و علیت	۱۱۳
۶-۵ ضریب هم‌بستگی رتبه‌ای اسپیرمن	۱۱۳
۷-۵ آزمون فرض درباره P	۱۱۵
خلاصه فصل پنجم	۱۱۸
خودآزمایی تشریحی فصل پنجم	۱۱۸
فصل ششم. رگرسیون و اصول پیش‌بینی	۱۲۱
هدف کلی	۱۲۱
هدف‌های یادگیری	۱۲۱
مقدمه	۱۲۱
۱-۶ استفاده از نمودار پراکندگی برای نمایش الگوی ارتباط بین داده‌ها	۱۲۲
۲-۶ رگرسیون خطی	۱۲۶
۳-۶ رگرسیون دو متغیره	۱۲۷
۴-۶ برآورد واریانس خطا	۱۳۰
۵-۶ پیش‌بینی	۱۳۱
خلاصه فصل ششم	۱۳۲
خودآزمایی تشریحی فصل ششم	۱۳۲

پیوست‌ها.....	۱۳۵
پیوست ۱: آشنایی مقدماتی با نرم‌افزار SPSS.....	۱۳۵
مقدمه	۱۳۵
حل مسئله	۱۶۵
پیوست ۲: پاسخ خودآزمایی‌ها	۱۷۷
پیوست ۳: جداول آماری	۱۹۰
منابع	۱۹۳

پیشگفتار

علم آمار امروزه در زمره علوم و فنون جدیدی است که اخیراً به وجود آمده است. شیوه‌ها و روش‌های آن به صورت علمی بسط و تعمیم یافته و علمی است که در تمام رشته‌های علوم، اعم از نظری یا عملی مورد استفاده قرار می‌گیرد. رشته علم اطلاعات و دانش‌شناسی یکی از رشته‌هایی است که نیاز مبرم دارد از روش‌های آماری برای معرفی کتاب‌سنجی و علم‌سنجی استفاده کند. لذا با مطالعه سرفصل‌های رشته علم اطلاعات و دانش‌شناسی بر آن شدیم که کتاب پیش رو را به رشته تحریر درآوریم. بنابراین، سعی شده است با ارائه تئوری علم آمار به صورت مختصر، کاربرد آن با ذکر مثال‌های مرتبط با رشته علم اطلاعات و دانش‌شناسی آورده شود. در تجزیه و تحلیل داده‌ها از نرم‌افزار آماری SPSS استفاده شده است. برای آشنایی با نرم‌افزار آماری SPSS در ضمیمه، یک نسخه آن به همراه محاسبات آورده شده است.

این کتاب در شش فصل تنظیم شده است. که در ادامه، مطالب هر یک از فصول به اختصار توضیح داده می‌شود.

عنوان فصل اول «کلیات و مفاهیم اولیه» است. در این فصل، سعی بر این است که شرح مختصری از واژه کتاب‌سنجی ارائه و در بخش‌های بعدی، مفاهیم اولیه آمار توصیفی به همراه مقیاس‌های اندازه‌گیری آورده شود.

عنوان فصل دوم «طبقه‌بندی داده‌های آماری و محاسبه شاخص‌های آماری» است. در این فصل، جدول توزیع فراوانی، انواع نمودارهای آمار و نحوه محاسبه شاخص‌های مرکزی و غیرمرکزی ارائه شده است.

عنوان فصل سوم «متغیر تصادفی و توزیع آماری» است. در این فصل، پس از تعریف متغیر تصادفی، چند توزیع آماری از نوع گسسته و پیوسته آورده می‌شود و در پایان، توزیع‌های نرمال، نرمال استاندارد و توزیع استودنت معرفی می‌شوند.

عنوان فصل چهارم «آزمون فرضیه آماری» است. در این فصل پس از آشنایی با مفاهیم اولیه آزمون فرض، فرض‌های آماری درباره میانگین جامعه نرمال و نسبت جامعه ارائه و آزمون می‌شود.

عنوان فصل پنجم «همبستگی و انواع آن» است. در این فصل، ضمن تعریف همبستگی دو متغیر تصادفی، انواع ضرایب همبستگی مورد بحث قرار می‌گیرد.

عنوان فصل ششم «رگرسیون و اصول پیش‌بینی» است. در این فصل، پس از آشنایی با نمودار پراکنش و خط رگرسیون، معادله خط رگرسیون با استفاده از نمونه برازش و براساس آن متغیر پاسخ پیش‌بینی می‌شود.

امیدواریم این کتاب قابل استفاده خوانندگان باشد و مورد علاقه آنان قرار گیرد، و آرزو مندیم خوانندگان به‌ویژه دانشجویان، نظرها و پیشنهادهای خود را برای مؤلفان ارسال دارند تا در تجدیدنظر چاپ‌های بعدی مورد استفاده قرار گیرند.

در پایان از سرکار خانم فاطمه محمدی که حروف‌چینی و صفحه‌آرایی کتاب را بر عهده داشتند تشکر می‌کنیم.

دکتر پرویز نصیری

دکتر ثریا ضیائی

دکتر باقر مقدس‌زاده

فصل اول

کلیات و مفاهیم اولیه

هدف کلی

آشنایی با مفاهیم اولیه کتابداری و آمار توصیفی.

هدف‌های یادگیری

پس از مطالعه این فصل، شما باید بتوانید:

۱. با واژه کتاب‌سنجی آشنا شوید.
۲. با مهم‌ترین مشخصه‌های کتاب‌سنجی آشنا شوید.
۳. با مطالعات آماری بین‌المللی در حوزه‌های سنجشی آشنا شوید.
۴. با مفاهیم اولیه آمار توصیفی آشنا شوید.
۵. با مقیاس‌های اندازه‌گیری آشنا شوید.

مقدمه

کمپبل^۱ در سال ۱۸۹۶ با مطالعه پراکندگی موضوعی انتشارات به کمک روش‌های آماری، اقدام به معرفی کتاب‌سنجی^۲ کرد. بعد از ایشان تعداد زیادی از نویسندگان ضمن تأیید انواع مطالعات کتاب‌سنجی توسط کمپبل سعی کردند که با محاسبه تعداد مدارک علمی، چشم‌اندازی در تاریخ علم فناوری ایجاد کنند. واژه قدیمی‌تر کتاب‌سنجی، کتاب‌شناسی آماری بود که در سال ۱۹۴۸ رانگاتان^۳ واژه کتاب‌سنجی را به کاربرد و آن را ترکیبی از روش‌های کمی قابل استفاده برای مدیریت و خدمات کتابخانه معرفی کرد [نوروزی چالکی]. اکنون

1. Campbell
2. Bibliometrice
3. Rangatan

می‌توان ادعا کرد که نزدیک به چهار دهه است که از واژه کتاب‌سنجی استفاده می‌شود. هم‌اکنون مطالعات کتاب‌سنجی در سطح وسیع در کتابخانه‌ها و جامعه علمی و پژوهشی به‌منظور سنجش و ارزیابی انتشارات علمی استفاده و روزبه‌روز نیز بر دامنه فنون و روش‌های آن افزوده می‌شود. البته نباید از تأثیر قوانینی که در قرن بیستم مطرح شد و کتاب‌سنجی را توسعه داد غافل بود. در همین زمینه باید به قانون زیپف^۱ اشاره کرد که برای نخستین بار به مطالعه مشخصه‌ها و فرایندهای مرتبط با مدارک پرداخت که بر توسعه کتاب‌سنجی تأثیر به‌سزایی گذاشت. او به‌دلیل معرفی «اصل کمترین کوشش» خود، بسیار معروف شد. در واقع اصل کمترین کوشش به این معناست که یک شخص می‌کوشد مشکلات فوری و به احتمال، دشواری‌های آینده‌اش را از طریقی که به کمترین تلاش نیاز داشته باشد، حل کند. زیپف در سال ۱۹۴۹ از اصل کمترین کوشش برای توصیف نرخ متوسط کمترین کار احتمال استفاده کرد و بر همین مبنا، اهمیت کوتاه‌نویسی مقاله‌ها را از طریق به‌کارگیری واژه‌های کمتر برای مفهوم، مورد تأکید قرار داد. اصل زیپف که در ابتدا برای هدایت رفتار اطلاع‌یابی استفاده می‌شد، به‌طور غیرمستقیم به فرمول زیپف که از مدل قانون توان^۲ در زبان‌شناسی کمی برگرفته شده‌است ربط داده می‌شود؛ به‌طوری که در زبان‌شناسی کمی مدل فرضی قانون-توان به‌صورت زیر ارائه می‌شود:

$$rf = C$$

که در آن r رتبه هر واژه و f فراوانی رخداد هر واژه و C مقدار ثابت است. در واقع حاصل ضرب رتبه هر واژه و فراوانی رخداد آن، مقداری است ثابت. قابل ذکر است علاوه بر قانون زیپف می‌توان از قوانین لوتکا، بردفورد^۳ و قانون پارتو^۴ نیز در مقام زیربناهای فکری کتاب‌سنجی نام برد. قدر مسلم آنکه این قوانین به‌گونه‌ای با زبان ریاضی بیان شده‌اند که می‌توان آن‌ها را در محیط واقعی به‌کار گرفت. قابل ذکر است که قانون زیپف فراوانی مشاهده شده واژگان در یک متن را بررسی می‌کند؛ درحالی‌که قانون لوتکا^۵ بهره‌وری نویسندگان را برحسب انتشارات علمی‌شان مورد مطالعه قرار می‌دهد. درحالی‌که بردفورد، به مطالعه پراکندگی مقاله‌های مجله‌های گوناگون می‌پردازد.

1. Zipf
2. Power-Law Model
3. Bradford
4. Pareto's Rule
5. Lotka

۱-۱ حوزه فعالیت‌های کتاب‌سنجی و کاربرد آمار

در سال ۱۹۶۹ پریچارد^۱ نخستین تعریف رسمی را از واژه کتاب‌سنجی ارائه کرد که عبارت است از:

«کاربرد روش‌های ریاضی و آمار برای کتاب‌ها و دیگر رسانه‌های ارتباطی». سن‌گوپتا^۲ ریشه کتاب‌سنجی را از معادل فارسی دو واژه کتاب و سنجش می‌داند. ریشه واژه کتاب را از کلمه ببلیون^۳ که خود دارای ترکیبی لاتین-یونانی است و معادل واژه ببیل یا ببلس^۴ به معنی کتاب یا کاغذ است معرفی می‌کند. همچنین وی واژه سنجش را با مفهوم اندازه‌گیری یکی گرفته و براین باور است که ریشه آن به واژه یونانی «مترکِلوس»^۵ یا کلمه لاتین متریکیوس^۶ به معنای اندازه‌گیری بازمی‌گردد. در سال ۲۰۰۲ بورگمن و فرنر^۷ کتاب‌سنجی را به‌طور ساده، اندازه‌گیری مشخصه‌ها و فرایندهای مرتبط با مدارک تعریف می‌کنند که از جمله این مشخصه‌ها و فرایندهای مرتبط با مدارک، می‌توان به موارد زیر اشاره کرد:

الف) تحلیل بسامد یا فراوانی واژه‌ها

ب) تحلیل هم‌واژه‌ها

ج) تحلیل استنادی

د) محاسبه ساده مدارک (مثلاً تعداد آثار منتشرشده علمی یک نویسنده یا یک گروه پژوهشی یا یک کشور)

قابل ذکر است که کتاب‌سنجی از تحلیل استنادی استفاده می‌کند و به آن توجهی خاص نشان می‌دهد؛ زیرا همواره این باور وجود دارد که یک دانشمند به دنبال شناسایی و مطالعه مقاله‌های مرتبط است. بنابراین شناسایی اینکه کدام مقاله به یکدیگر استناد کرده‌اند، می‌تواند برای او کارساز باشد و به وی در شناسایی مرتبط‌ترین مقاله‌هایی که در یک حوزه وجود دارند، کمک کند. از همین رو است که نمایه‌های استنادی^۸ در مقام ابزاری قدرتمند برای انجام تحلیل‌های استنادی در کتاب‌سنجی اهمیتی ویژه دارد. در واقع، نمایه‌های استنادی علاوه بر بهره‌گیری از تفکر تحلیل استنادی، تا حد زیادی با

1. Pritchard
2. San Gupta
3. Biblion
4. Biblos
5. Metrokos
6. Metricus
7. Borgman and Furner
8. Citation Indexes

قانون پراکندگی بردفورد نیز سازگارند. افزون بر استفاده از تحلیل‌های استنادی، استفاده از شاخص‌ها و فنون دیگر نیز ضرورت دارد تا بتواند نتایج متنوع‌تر، معنادارتر و مفیدتر را برای سیاست علم فراهم کند. با این وجود، امروزه نیز به‌طور دائم، فنون مکمل برای کتاب‌سنجی پا به عرصه وجود می‌گذارند. انقلابی که با ورود وب ایجاد شده، بر این امر دامن زد و کتاب‌سنجی را نیز تحت تأثیر خود قرار داد. استفاده از شاخص‌های مبتنی بر استناد، هنوز هم هسته اصلی این نوع مطالعات را تشکیل می‌دهد. گذشته از روش‌های اصلی تحلیل استنادی، بزرگ‌ترین تحول در کتاب‌سنجی از محیط ارتباطات علمی و دسترس‌پذیری به منابع اطلاعاتی جدید و مهمی نظیر پروانه‌های ثبت اختراعات، صفحه‌های وب و آمارهای استفاده از کتابخانه‌ها نشأت گرفته است.

البته باید توجه داشت که کتاب‌سنجی، هرگز تنها به مقاله‌های علمی وابسته نبوده و سایر منابع علمی نیز همواره مورد توجه پژوهشگران این حوزه قرار داشته‌اند.

۱-۲ مهم‌ترین مشخصه‌های کتاب‌سنجی

مطالعات کتاب‌سنجی با هدف توسعه خدمات کتابخانه‌ها و مراکز اطلاع‌رسانی انجام می‌شوند. در حقیقت، کتابداران همیشه فنون کتاب‌سنجی را در امور گوناگون کتابخانه‌ای به کار می‌گرفتند و از دیرباز در امور کتابخانه‌ای رایج بوده است. کتاب‌سنجی سرمنشأ بسیاری از حوزه‌های سنجشی از جمله علم‌سنجی، اطلاع‌سنجی و وب‌سنجی است. بنابراین، کتاب‌سنجی از نمایه‌های استنادی در جایگاه یک ابزار استفاده می‌کند و به‌طور مستقیم روش‌های ریاضی و آمار را برای دستیابی به هدف‌های خود به کار می‌برد. به این ترتیب، کتاب‌سنجی حوزه‌ای است به‌طور کامل کمی که با تکیه بر روش‌های ریاضی و آمار، به سنجش و ارزیابی متون می‌پردازد.

۱-۳ کتاب‌سنجی نوین در ایران

در ایران عده‌ای از صاحب‌نظران علم اطلاعات و دانش‌شناسی، توجه به بنیان‌های تاریخی و مفهومی کتاب‌سنجی را در اوایل دهه ۱۳۶۰ (مطابق ۱۹۸۰ میلادی) مطرح کردند که می‌توان به مقالات دیانی (۱۳۶۱)، حری (۱۳۶۲)، دیانی و عصار (۱۳۶۷)، مهراد (۱۳۷۲) و عصار (۱۳۷۶ الف، ۱۳۷۶ ب و ۱۳۷۷) اشاره کرد. در سال ۱۳۸۹، منصوریان با هدف شناسایی مهم‌ترین تحقیقات کتاب‌سنجی و علم‌سنجی در ایران،

کلیات و مفاهیم اولیه ۵

بدون جدایی این دو حوزه و با شناسایی چند محور پژوهشی فرعی، کتاب‌سنجی و علم‌سنجی در ایران را در چهار گروه کلی زیر دسته‌بندی کرد:

- مطالعه ارتباطات علمی و تحلیل استنادی
 - ارزیابی کمی و کیفی منابع و انتشارات علمی
 - سنجش برون‌داد، بازدهی و تأثیرگذاری علمی
 - تدوین سیاست‌ها و خط‌مشی‌های علمی و پژوهشی
- به‌لحاظ تاریخی، واژه اطلاع‌سنجی^۱ در مقایسه با حوزه‌های کتاب‌سنجی و علم‌سنجی، قدمتی کمتر دارد و اولین بار در سال ۱۹۷۹ در کشور آلمان از این واژه استفاده شد.

ناکه^۲، از اطلاع‌سنجی، برای پوشش‌دادن آن بخش از علم اطلاعات که با اندازه‌گیری پدیده اطلاعات و کاربرد روش‌های ریاضی برای حل مسائل رشته‌ها سروکار دارد، استفاده کرد و آن را در حوزه‌هایی مانند کتاب‌سنجی، در بخش‌هایی از نظریه بازیابی اطلاعات و حتی شاید در سطحی بسیار وسیع‌تر قابل استفاده دانست (نقل در: هود و ویلسون، ۲۰۰۱: ۲۹۴).

ناکه در سال ۱۹۸۴ و در شرایطی که هنوز اطلاع‌سنجی، واژه‌ای برگرفته از کتاب‌سنجی و علم‌سنجی محسوب می‌شد، «مؤسسه مشترک اطلاعات علمی و فنی (وینیتی)»^۳، «کمیته اطلاع‌سنجی»^۴ را در «فدراسیون بین‌المللی دبیزش (فی)»^۵، تأسیس کرد. بروکس^۶ (۱۹۸۸) در «نخستین کنفرانس بین‌المللی کتاب‌سنجی و جنبه‌های نظری بازیابی اطلاعات»^۷، اطلاع‌سنجی را به‌لحاظ مفهومی به دو حوزه کتاب‌سنجی و علم‌سنجی تقسیم کرد و آن را برای ارزیابی هر دو نوع اطلاعات الکترونیکی و اطلاعات سنتی قابل استفاده دانست؛ با وجود این، در اوایل دهه ۱۹۹۰ میلادی، مطالعات اطلاع‌سنجی، گستردگی بیشتری یافت. بروکس (۱۹۹۰) در «دومین کنفرانس بین‌المللی کتاب‌سنجی و جنبه‌های نظری بازیابی اطلاعات»^۸، باز هم، اطلاع‌سنجی را اصطلاحی اعم از علم‌سنجی و کتاب‌سنجی معرفی و درعین‌حال، بر این نکته تأکید کرد که علم‌سنجی بیشتر به‌سوی

1. Informetrics

2. Nacke

3. All-union Institute for Scientific and Technical Information (VINITI)

4. Committee of Informetrics

5. Federation Internationale de le Documentation (FID)

6. Brookes

7. First International Conference on Bibliometrics and Theoretical Aspects of Information Retrieval

8. Second International Conference on Bibliometrics and Theoretical Aspects of Information Retrieval

مطالعات سیاستی گرایش دارد و حال آنکه کتاب‌سنجی بیشتر در راستای کاربردهای کتابخانه‌ای استفاده می‌شود؛ علاوه بر آن، تگ- ساتکلیف^۱ (۱۹۹۲) نیز در همان کنفرانس، حوزه اطلاع‌سنجی را یکی از حوزه‌های مستخرج از کتاب‌سنجی و علم‌سنجی معرفی کرد. با این حال، در سومین مجموعه مقالات آن کنفرانس که در سال ۱۹۹۲ (با عنوان سومین کنفرانس بین‌المللی اطلاع‌سنجی)^۲ برگزار شد، جایگاه اطلاع‌سنجی، استحکام و قوامی بیشتر یافت؛ هر چند، در «چهارمین کنفرانس بین‌المللی کتاب‌سنجی، اطلاع‌سنجی و علم‌سنجی»^۳ تعداد کمتری مقاله در زمینه اطلاع‌سنجی ارائه شد و بیشتر مقاله‌ها به حوزه‌های علم‌سنجی و کتاب‌سنجی اختصاص داشتند؛ در این میان، سه جلد از چهار جلد مجموعه مقالات چهارمین کنفرانس بین‌المللی کتاب‌سنجی، اطلاع‌سنجی و علم‌سنجی، به‌طور کامل به مسائل مجله‌های انگلیسی‌زبان اختصاص یافت (گلنزل؛ کرتشمر، ۱۹۹۴ab، ۱۹۹۲).^۴ در سال ۱۹۹۲ و در همان کنفرانس، «انجمن بین‌المللی علم‌سنجی و اطلاع‌سنجی (آی.اس.اس.آی)»^۵ تأسیس شد؛ این انجمن از آن پس، هر دو سال یک‌بار «کنفرانس بین‌المللی آی.اس.اس.آی» را برگزار کرد؛ علاوه بر آن، ظهور مجله «مدیریت و پردازش اطلاعات»^۶ در سال ۱۹۹۲ را باید یکی دیگر از مهم‌ترین حرکتهایی دانست که در حوزه اطلاع‌سنجی به انجام رسید. در مجموع، باید یادآوری شود که اوایل دهه ۱۹۹۰، اصطلاح اطلاع‌سنجی در سطحی گسترده‌تر شناخته شد و پس از آن تداوم یافت (نقل در: هود و ویلسون، ۲۰۰۱: ۲۹۴).

اطلاع‌سنجی اغلب با تولید اطلاعات یا فرایندهای تولید اطلاعات^۷ سروکار دارد. آن‌ها مصداق‌های اصلی اطلاعات را «منبع»^۸ و «فقره اطلاعاتی»^۹ نامیدند و برای نمونه، مجله را به‌منزله «منبع» و مقاله را به‌منزله «فقره‌های اطلاعاتی» در نظر گرفتند؛ بنابراین از نظر آن‌ها، اطلاع‌سنجی با اثبات روابط میان منابع و فقره‌های اطلاعاتی و توصیف قواعد موجود در الگوهای تولید اطلاعات سروکار دارد. در این میان، دایادتو^{۱۰} (۱۹۹۴) مباحث موجود در اطلاع‌سنجی را به سه قسمت تقسیم کرده است:

1. Tague- Sutcliffe
2. Third International Conference on Informetrics
3. International Conference on Bibliometrics, Informetrics and Scientometrics
4. GlÄnz, Kretschmer
5. Institute for Scientific Information
6. Information Processing & Management
7. Information Production Processes (IPPs)
8. Source
9. Item Information
10. Dayadto

- توزیع‌های کتاب‌سنجی / اطلاع‌سنجی
- تحلیل‌های استنادی
- شاخص‌های اطلاع‌سنجی

از سویی دیگر، إرار (۲۰۰۰)^۱ به‌کارگیری روش‌های آماری پیشرفته و برخی روش‌های پژوهش عملیاتی و ریاضی را در حوزه علوم علم اطلاعات و دانش‌شناسی، به‌عنوان اطلاع‌سنجی معرفی می‌کند.

۱-۴ مطالعات بین‌المللی در حوزه‌های سنجشی

برای مشخص‌کردن وضعیت کنونی و گذشته کتاب‌سنجی و علم‌سنجی از دیدگاه‌های مختلف، جست‌وجو در پایگاه‌های اطلاعاتی سه‌گانه آ.اس.ای تحت SCI، SSCI، HCI و H از طریق وب‌سایت علوم (Web of Science) امکان‌پذیر است؛ به‌طور نمونه می‌توان به مؤلفان، کشورها، مؤسسه‌ها، زبان‌ها، منبع و موارد دیگر اشاره کرد. از آنجایی که رکوردها براساس پارامترهای مختلف از جمله مؤلفان، کشورها، مؤسسه‌ها، زبان‌ها، منبع و موارد دیگر در پایگاه‌ها تعبیه شده‌اند، می‌توان با استفاده از نرم‌افزار تحلیلی، اطلاعات را مورد تجزیه و تحلیل قرار داد. برای نمونه می‌توان ده نویسنده پرکار، ده کشور برتر، ده دانشگاه یا مؤسسه برتر، ده زبان مورد استفاده و یا ده مجله برتر در تولید مدارک کتاب‌سنجی را به‌صورت گزارش تهیه کرد.

جدول ۱-۱. جدول توزیع فراوانی ده نویسنده پرکار در تولید مدارک کتاب‌سنجی

ردیف	نام نویسندگان	تعداد
۱	GALANZEL, W	۴۱
۲	MOED, HF	۳۰
۳	KOSTOFF, RN	۴۰
۴	VAN RAAN, AFJ	۲۱
۵	LEWISON, G	۱۹
۶	BORDONS, M	۱۸
۷	THELWAL, M	۱۸
۸	VAN RAAN, AFJ	۱۸
۹	GOMEZ, I	۱۵
۱۰	NEDEFHOF, AJ	۱۵

جدول ۱-۲. جدول توزیع فراوانی ده کشور برتر در تولید مدارک کتاب‌سنجی

ردیف	نام کشورها	تعداد مدارک
۱	ایالات متحده آمریکا	۴۵۳
۲	اسپانیا	۲۲۷
۳	انگلستان	۱۵۶
۴	هلند	۱۳۶
۵	آلمان	۱۱۷
۶	ایتالیا	۷۱
۷	کانادا	۷۱
۸	تایوان	۶۷
۹	فرانسه	۶۷
۱۰	بلژیک	۶۷

جدول ۱-۳. جدول توزیع فراوانی ده دانشگاه یا مؤسسه برتر در تولید مدارک کتاب‌سنجی

ردیف	نام دانشگاه‌ها یا سازمان‌ها	تعداد مدارک
۱	LEIDEN UNIV	۷۸
۲	CSIC	۴۳
۳	INDIANA UNIV	۳۴
۴	UNIV GRANADA	۳۱
۵	HUNGARIAN ACAD SCI	۲۸
۶	OFF NAVAL RES	۲۸
۷	KATHOLIEKE UNIV LEUVEN	۲۵
۸	UNIV SUSSEX	۲۵
۹	CITY UNIV LONDON	۲۳
۱۰	UNIV VALENCIA	۲۲

جدول ۱-۴. جدول توزیع فراوانی زبان‌های مورد استفاده در تولید مدارک کتاب‌سنجی

ردیف	زبان	تعداد مدارک
۱	انگلیسی	۱۵۰۳
۲	اسپانیایی	۱۲۱
۳	آلمانی	۴۵
۴	فرانسوی	۱۹
۵	روسی	۸
۶	ایتالیایی	۳
۷	ژاپنی	۳
۸	کرواتی	۲
۹	قزاقستانی	۲
۱۰	صربی-کروانی	۲
۱۱	پرتغالی	۲

۹ کلیات و مفاهیم اولیه

جدول ۱-۵. جدول توزیع فراوانی مجلات چاپ‌کننده مدارک مرتبط با کتاب‌سنجی

ردیف	عنوان مجلات	تعداد
۱	SCIENTOMETRIC	۱۱۱
۲	JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY	۲۸
۳	RESEARCH EVALUATION	۱۳
۴	INFORMATION PROCESSING & MANAGEMENT	۱۲
۵	TECHNOLOGICAL FORECASTING AND SOCIAL CHANGE	۱۱
۶	SOCIAL WORK IN HEALTH CARE	۹
۷	JOURNAL OF DOCUMENTATION	۸
۸	JOURNAL OF INFORMATION SCIENCE	۸
۹	TECHNOVATION	۷
۱۰	EUROPEAN JOURNAL OF PUBLIC HEALTH	۶

۱-۵ مفاهیم اولیه آمار توصیفی

آمار، علمی است که مشخصات یا خصوصیات جامعه‌های آماری را با توجه به شرایط کیفی مربوط، مورد مطالعه قرار می‌دهد. برای ارتقای توانایی در توصیف و تفسیر انبوهی از اطلاعات پژوهش‌های علوم مختلف، به آمار نیازمندیم. آمار توصیفی که شاخه‌ای از آمار است، ابزاری برای توصیف انبوهی از داده‌هاست. گردآوری اطلاعات و ارائه روش دقیق آن‌ها کمک مهمی در فهم و تفسیر صحیح آن‌ها است. برای ارائه مشاهدات کمی معمولاً از دو روش استفاده می‌شود. یک روش، شامل نشان دادن خلاصه‌شده‌ای از اطلاعات است که معمولاً در قالب جداول آماری ارائه می‌شود. روش دیگر، مشاهدات را در قالب تصویر به صورت نمودارها و یا دیگر شکل‌های مشابه نشان می‌دهد. لذا در این بخش، پس از ارائه تعریفی از جامعه آماری و نمونه، روش‌های نمونه‌گیری، انواع متغیرها و مقیاس‌های اندازه‌گیری شرح داده می‌شود.

۱-۵-۱ جامعه^۱

جامعه عبارت است از گروهی از افراد، اشیاء یا حوادث که حداقل یک صفت یا ویژگی مشترک دارند و به خاطر همان صفت، مورد بررسی قرار می‌گیرند. مفهوم جامعه از نظر آماری خیلی وسیع‌تر از مفهوم واژه آن است. مانند ساکنان کلان‌شهرها، کلیه کتب در کتابخانه‌های کشور و یا دانشجویان دانشگاه‌ها. برای انجام هر کار آماری روی یک

1. Population

جامعه باید آن جامعه و ویژگی مورد مطالعه، قبلاً بدون هرگونه ابهام مشخص شوند. به عنوان مثال، جامعه دانشجویان دانشگاه پیام‌نور مشهد و میزان استفاده از کتابخانه. جامعه آماری ممکن است، با پایان (محدود) باشد. یعنی قادر به شمارش اعضای آن جامعه باشیم و یا بی‌پایان (نامحدود) باشد که در این صورت قادر به شمارش اعضای آن جامعه نیستیم. مثلاً دانشجویانی که طی دو سال گذشته در دانشگاه پیام‌نور ثبت‌نام کرده‌اند با پایان است، ولی جامعه افرادی که از مهرماه سال آینده به بعد در دانشگاه پیام‌نور ثبت‌نام خواهند کرد، بی‌پایان است. نمایش اعضای جامعه برای گرفتن نمونه و یا انجام چهار عمل اصلی بین آن‌ها نقش مهمی در استنباط آماری دارد. در حالتی که جامعه آماری محدود یا متناهی و دارای حجم N باشد، اعضای جامعه را به صورت زیر نمایش می‌دهند:

$$X_1, X_2, X_3, \dots, X_N$$

که در آن X_1 عضو اول، X_2 عضو دوم و X_i عضو i ام جامعه است. برای مثال اگر اعضای انجمن علم اطلاعات و دانش‌شناسی در کشور دارای ۱۰ عضو باشد، می‌توان بدون مراجعه به آن‌ها، آن‌ها را به صورت زیر نمایش داده و یا در ذهن شبیه‌سازی کرد.

$$X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}$$

البته به این نکته باید اشاره کرد که X_1 ، نفر اول انجمن نیست، بلکه ویژگی‌ای است که می‌تواند برای یکی از ۱۰ نفر انجمن تخصیص داده شود. در حالت کلی X_i می‌تواند، سابقه کار، میزان درآمد، میزان تحصیلات، سن و غیره برای نفر i ام باشد. بنابراین به اطلاعاتی که با روش‌های مختلف به دست می‌آید، داده آماری گفته می‌شود که در واقع، همان اطلاعات خامی است که هنوز هیچ‌گونه بررسی‌ای روی آن‌ها انجام نشده است.

گاهی جامعه مورد مطالعه به قدری انبوه است که عملاً می‌توان آن را نامحدود تلقی کرد؛ مثل جامعه کلیه کتب منتشرشده در بازار نشر ایران طی ده سال گذشته. از آنجایی که اغلب، مطالعه یکایک افراد جامعه به علت هزینه زیاد، کمی وقت، نداشتن امکانات کافی و یا از بین رفتن جامعه بر اثر بررسی، مقدور نیست، قسمتی از جامعه را به جای تمام آن در نظر می‌گیرند و به این بخش در نظر گرفته شده، نمونه می‌گوییم که در بخش بعدی ارائه می‌شود.

۱-۵-۲ نمونه^۱

قسمتی از جامعه را که طبق ضوابطی مقبول انتخاب و مطالعه آن به جای مطالعه تمام جامعه مقدور است، نمونه‌ای از جامعه می‌نامیم. معمولاً به مصداق مشت نمونه خروار است نتیجه حاصل از مطالعه نمونه را به تمام جامعه تعمیم می‌دهند. ولی این کار احتیاط دارد؛ زیرا هر مشت نمی‌تواند نمونه خروار باشد و قطعاً بی‌غرضی در انتخاب مشت و اندازه مشت در این نمایندگی نقش مهمی دارد. مسئله انتخاب یک نمونه خوب به قدری مهم است که قسمت زیادی از نظریه‌های آمار و احتمال به آن اختصاص دارد. اغلب برای کسب بی‌طرفی، نمونه‌هایی در نظر گرفته می‌شود که انتخاب آن‌ها کاملاً شانسی باشد. به این نوع از نمونه‌ها، نمونه تصادفی می‌گوییم. اگر حجم نمونه n باشد، می‌توان اعضای نمونه را که بخشی از جامعه است به صورت زیر نمایش داد:

$$X_1, X_2, X_3, \dots, X_n$$

که در آن X_1 عضو اول، X_2 عضو دوم و X_i عضو i ام نمونه است. باید متذکر شد که X_1 که اولین نمونه است، اولین عضو جامعه نیست و می‌تواند هر کدام از $X_1, X_2, X_3, \dots, X_N$ باشد $n \leq N$. همان‌طور که گفته شد، انتخاب نمونه نقش مهمی در استنباط دارد که در ادامه، انواع روش‌های نمونه‌گیری ارائه می‌شود.

۱-۵-۳ روش‌های نمونه‌گیری

لازم به توضیح است برای داشتن نمونه بهتر، می‌توان از روش‌های نمونه‌گیری استفاده کرد که مهم‌ترین آن‌ها عبارت‌اند از:

- نمونه‌گیری تصادفی ساده
- نمونه‌گیری طبقه‌ای
- نمونه‌گیری منظم
- نمونه‌گیری خوشه‌ای

در ادامه، شرح مختصری از روش‌های نمونه‌گیری ارائه می‌شود.

- در نمونه‌گیری تصادفی ساده، درحالی‌که حجم جامعه متناهی است، انتخاب نمونه به روش نمونه‌گیری تصادفی ساده با جایگذاری و بدون جایگذاری انجام می‌شود