**نام مقاله: قانون برادفورد در مجموعه‌هاي نامعين/استنتاج‌هاي آماري براي تحليل كتابخانه**

**نام نشريه: فصلنامه كتابداري و اطلاع رساني (اين نشريه در www.isc.gov.ir نمايه مي شود)**

**شماره نشريه: 26 \_ شماره دوم، جلد 7**

**پديدآور: استفان جي بنسمن؛ ترجمة حيدر مختاري، آنسه حسيني‌زاده**

**چكيده**

**در اين نوشته زيربناي فكري و جنبه‌هاي عملي قانون برادفورد تشريح شده است. تفاوت مقوله‌بندي سن‍ّتي با مجموعه‌هاي نامعين «لطفي‌زاده» شرح داده شده، انديشه‌هاي «كارل پيرسون» دربارة «مسئلة بنيادي علم» آورده شده، به «داده‌هاي پرت» از ديد «بارنت» و «لوئيس» پرداخته شده و با تكيه بر اين موارد، گزارشي از يك بررسي كه در دانشگاه لويزيانا در پيوند با انتخاب مجلات شيمي توسط اعضاي هيئت علمي و مقايسه آن با «آي‌اس‌اي» انجام شد، ارائه گرديده و نتيجه گرفته شده است كه مجموعه‌هاي برادفورد نيز نامعين هستند و سرانجام چند شيوه براي مهاركردن «داده‌هاي پرت» آورده شده و اعلام شده است كه نظرية احتمالات و نظرية مجموعه‌هاي نامعين متناقض نيستند، اما قالب‌هايي مكمل براي تجزيه و تحليل به حساب مي‌آيند.**

**كليدواژه‌ها: قانون برادفورد، مجموعه‌هاي نامعين، تحليل‌هاي آماري، گزينش نشريه‌هاي ادواري**

از وقتي كه «اس.سي.برادفورد»*[[3]](http://www.aqlibrary.org/index.php?module=TWArticles&file=index&func=view_pubarticles&did=15&pid=10" \l "ftn3)* اولين مقالة خود را در حوزة قانون كتابسنجي منتشر ساخت، بيش از شصت سال مي‌گذرد. از آن زمان تا به حال دربارة قانون پراكندگي برادفورد بسيار نوشته‌اند؛ اما هنوز ناگفته‌هاي بسياري وجود دارد. من پس از انجام تحقيقات نظري و عملي دامنه‌دار در ارزيابي و كاربرد مجلات علمي، چنين نتيجه‌گيري كرده‌ام كه قانون برادفورد اساساً ملغمه‌اي از دو مفهوم بسيار مهم در كاربرد روش‌هاي آماري در امر تجزيه و تحليل كتابخانه ـ يعني توزيع احتمال و مجموعه‌هاي نامعين ـ است. درواقع، اين قانون را مي‌توان توصيف رياضي الگوي احتمال براي تشكيل مجموعه‌هاي نامعين دانست. در مقالة حاضر به استنتاج‌هاي آماري قانون برادفورد در ايجاد مجموعه‌هاي نامعين پرداخته شده است.

**قانون برادفورد**

«برادفورد» كتابدار ارشد كتابخانة موزة علوم در كنسينگتون*[[4]](http://www.aqlibrary.org/index.php?module=TWArticles&file=index&func=view_pubarticles&did=15&pid=10" \l "ftn4)* جنوبي لندن بود. هدف عمدة وي از اين تحقيق كه منجر به ايجاد قانون پراكندگي شد، بهينه‌سازي ميزان پوشش نشريات علمي در خدمات نمايه‌سازي و چكيده‌نويسي بود. نقص در اين پوشش كه با برآورد و مشاهدة حدود 000,500 مقالة علمي منعكس‌نشده در خدمات نمايه‌سازي و چكيده‌نويسي از مجموع 000,750 مقالة علمي منتشرشده در يك سال آشكار شد، «برادفورد» را به انديشه واداشت. علت اين امر شايد در شيوة پراكندگي و توزيع مطالب مربوط به يك موضوع معين در نشريات ادواري حاوي آن موضوع باشد. «برادفورد» فرضيه‌اي را از طريق اين تحقيق به صورت زير مطرح كرد:

... فرضية ديگر آن است كه توزيع فراواني ارجاعات در تمامي نشريات ادواري تا حد زيادي با حيطة موضوعي آن‌ها نسبت معكوس دارد. برطبق اين فرضيه، از لحاظ ارتباط محدودة موضوعي با موضوع موردنظر، مجموع نشريات ادواري را مي‌توان به مجموعه‌هايي تقسيم كرد؛ به نحوي كه در اين مجموعه، رده‌هاي دورتر به اندازة رده‌هاي مرتبط [با موضوع معين] در توليد مراجع نقش دارند. محدودة كل نشريات ادواري، بر اين اساس، به مثابه خانواده‌اي از نسل‌هاي متناوب هستند كه نسبت خويشاوندي آن‌‌ها پيوسته كم مي‌شود و هر نسل از نسل قبلي خود بيشتر است و برحسب درجة دوري، توليد نسل‌ها تنزل مي‌يابد.

اين فرضيه با دو دسته از ارجاعات برگرفته از كتابشناسي‌هاي جاري (گردآوري شده در كتابخانة موزه علوم*[[5]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn5%22%20%5Co%20%22)*) كه يكي مربوط به ژئوفيزيك كاربردي و ديگري مربوط به روغنكاري بود، مورد آزمون قرار گرفت. با توجه به اين‌كه نتايج حاصل از بررسي ارجاعات ژئوفيزيك كاربردي و روغنكاري يكسان بوده‌اند، در مقالة حاضر، تحليل خود را به ژئوفيزيك كاربردي محدود كرده‌ام. مجموعة ژئوفيزيك كاربردي شامل كل ارجاعات چهار سال (از 1928 تا 1931)، جمعاً 1332 ارجاع موجود در 326 عنوان مجله بوده است. در صورت مرتب‌شدن اين مجلات به صورت نزولي، رتبه‌بندي از يك مجله حاوي 93 ارجاع تا 109 مجله حاوي يك ارجاع در نوسان مي‌باشد. «برادفورد» در گزارش خود، نتايج حاصل از گروه‌بندي مجلات در هر دسته را بدين‌گونه در سه ردة زير جاي داد:

a. مجلاتي كه سالانه بيش از 4 ارجاع (به‌طور متوسط) داشته‌اند؛

b. مجلاتي كه سالانه حداكثر 4 و حداقل بيشتر از يك ارجاع داشته‌اند؛

c. مجلاتي كه سالانه داراي يك ارجاع يا كمتر بوده‌اند.

جدول 1 نشان‌دهندة گروه‌بندي «برادفورد» از مجلات ژئوفيزيك كاربردي و تعيين رده براي آن‌ها است. پرواضح است كه اين نتايج بر فرضيات آزمون‌شده مهر تأييد مي‌زنند؛ بنابراين، ردة a با وجود شمول بر 8/2 درصد مجلات، 2/32 درصد ارجاعات را دربرمي‌گيرد؛ ردة b با شمول بر 1/18 درصد مجلات، 5/37 درصد ارجاعات را به خود اختصاص داده است و بالاخره ردة c با دارابودن 1/79 درصد مجلات، فقط 3/30 درصد ارجاعات را دارد.

جدول 1 به وضوح نشان مي‌دهد كه با وجود افزايش نمايي تعداد مجلات در هر رده، تعداد ارجاعات موجود در آن‌ها تقريباً ثابت مي‌ماند. چشم‌انداز به‌دست‌آمده از اين داده‌ها، برادفورد را به ارائة فرمول‌بندي شفاهي زير از قانون پراكندگي رهنمون شد:

... در صورت رتبه‌بندي نزولي مجلات علمي برحسب ميزان كاهش توليد مقالات در يك زمينة معين موضوعي، مي‌‌توان آن‌ها را به نشريات ادواري هسته كه اختصاص بيشتري به موضوع موردنظر دارند و چند گروه يا ناحية حاوي همان تعداد مقالة موجود در گروه هسته تقسيم كرد؛ به طوري كه نشريات ادواري هسته و گروه‌هاي بعدي داراي نسبت 1:n:n2 باشد.

البته برادفورد ارتباط تعداد مقالات در موضوع معين از مجلات مختلف را با اندازة آن‌ها بررسي نكرده است. چون دغدغة عمدة وي همسوساختن كارگزاران خدمات نمايه‌سازي و چكيده‌نويسي از طريق ايجاد راهكارهايي براي اطمينان از دسترسي كامل آن‌ها به تمام عناوين موضوعي بوده، به اين ارتباط توجهي نشان نداد. توجه «برادفورد» به مقالات يك موضوع معين و طرز توزيع آن‌ها در عناوين مختلف بود و بنابراين بحث اصلي ما درصد مقالات است. براي دسته‌بندي مجلات به شيوة استاندارد، بايد موضوع را از ديد خود مجلات ـ نه از ديد مقالات ـ بررسي كرد و ميزان درصد مقالات در يك مجلة خاص مربوط به يك موضوع خاص را تعيين نمود. مي‌توان چنين اظهارنظر كرد كه عنواني با بيشترين مقاله در يك موضوع معين‌ ـ ‍‌بويژه يك موضوع اخص ـ يك مجلة چندرشته‌اي با نسبت كمي از مقالات مربوط به آن موضوع است. من در اينجا از روش دسته‌بندي برادفورد برحسب درصد مقالات مربوط به موضوع ـ و نه برحسب درصد مقالات در يك مجلة معين ـ پيروي كرده‌ام. با اين‌حال، اين فرض برگرفته از قانون برادفورد كه «نشريات ادواري هسته بيشترين اختصاص را به موضوع موردنظر دارند» درواقع نشانگر مجلات پاية حاوي بيشترين مقاله در يك موضوع معين است و بناي مقالة حاضر بر اين فرض نهاده شده است.

از نظرگاه «تئوري مجموعه‌هاي نامعين» جنبة مهم قانون پراكندگي برادفورد، ترسيم واقعيت مربوط به نقطة نهايي در فرضية اوليه مي‌باشد، مبني بر اينكه در هر مجموعة موضوعي معين «محدودة كلي نشريات ادواري ... به صورت خانواده‌هايي از نسل‌هاي متوالي است كه در آن، درجة خويشاوندي پيوسته كم و كمتر مي‌شود، تعداد هر نسل از نسل قبلي آن بيشتر است، و هر نسل بسته به ميزان دوري و كم‌شدن ميزان خويشاوندي، توليد و زايش كم و كمتري دارد».

**جدول 1. رده‌بندي مجلات از ديدگاه برادفورد در ژئوفيزيك كاربردي از سال 1928 تا 1931**

|  |  |  |
| --- | --- | --- |
| **رده** | **مجلات** | **ارجاعات** |
| **تعداد** | **درصد** | **تعداد** | **درصد** |
| **a**. بيش از چهار ارجاع در سال [93ـ17 ارجاع] | 9 | 8/2 | 429 | 2/32 |
| **b**. 4ـ2 ارجاع در سال [16ـ5 ارجاع] | 59 | 1/18 | 499 | 5/37 |
| **c**. 1ـ1/0 ارجاع در سال [4 تا ارجاع] | 258 | 1/79 | 404 | 3/30 |
| كل | 326 | 100 | 1332 | 100 |

**مجموعه‌هاي نامعين*****[[6]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn6%22%20%5Co%20%22)* از ديدگاه «لطفي‌زاده»**

تئوري كلاسيك مجموعه‌ها برپاية اين انديشه قرار دارد كه مي‌توان بين گروه‌ها تفاوت‌هاي جزئي و مشخص قائل شد. بر اساس اين تئوري، امكان تعيين عضويت/ عدم عضويت قطعي يك چيز در يك گروه وجود دارد. «كوسكو»*[[7]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn7%22%20%5Co%20%22)* به نقل از «ارسطو»*[[8]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn8%22%20%5Co%20%22)*، ايده‌هاي وي را از اين مفهوم در كتابش درباب منطق نامعين اين‌گونه بيان مي‌كند:

منطق دودويي ارسطو به يك قانون تعميم‌پذيري دارد: A يا نه A، يا اين يا نه اين. آسمان يا آبي است، يا آبي نيست و نمي‌تواند هم آبي باشد و هم آبي نباشد، «هم A هم نه A» امكان ندارد.

«A يا نه A» بياني ساده از قانون ميانگين خروج از عضويت است. در عملكردهاي مجموعة كلاسيك، به A عدد يك و به نه A عدد صفر را تخصيص مي‌دهند.

«برادفورد» بي‌شك بر اساس تئوري مجموعة كلاسيك مي‌انديشيده است. اين رويكرد از مقاله‌اي كه وي در «انجمن بريتانيايي كتابشناسي بين‌المللي»*[[9]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn9%22%20%5Co%20%22)* در سال 1944 ارائه داد، كاملاً برمي‌آيد. وي در اين مقاله با كاوش در بنيادهاي رده‌بندي دهدهي جهاني، رده را «مجموعه‌اي از موجوديت‌ها و اشيا با يك ويژگي مشترك» مي‌داند. وي از نظام نمادها و منطق مطرح‌شده در كتاب «جورج بول»*[[10]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn10%22%20%5Co%20%22)* با عنوان **قوانين انديشه** (1854) كه منطق شكل‌گيري چنين مجموعه‌هايي در ذهن و انديشة آدمي را تبيين مي‌كند، استفاده كرد. وي با به‌دست آوردن معادله‌اي كه مشتق از بنيادهاي حساب منطقي بول بود، به نتيجة زير رسيد:

بنابراين به قانوني دست مي‌يابيم كه برطبق آن «هم تعلق و هم عدم تعلق يك ويژگي به چيزي واحد محال است». اين همان اصل تعارض ارسطو مي‌باشد كه [بول] آن را به عنوان اصلي بسيار متقن و مبرهن، باز شناخت.

اين‌كه چنين معادله‌اي از درجة دو با دو ريشة صفر و يك است، نشان مي‌دهد كه ما عمل طبقه‌بندي را با انتزاع به جفت‌هاي متضاد، مثل مردان و غيرمردان انجام مي‌دهيم و درمي‌يابيم كه براي هر ردة مشخص شده با نماد فرضي x، فقط ارزش صفر و يك مي‌توان درنظر گرفت.

با وجود اين «برادفورد» از اين امر كاملاً آگاهي داشت كه «مانعةالجمع نبودن رده‌ها، همواره عملي و ممكن نيست».

در نوشته‌ها، به منظور مشخص شدن تمايز بين مجموعه‌هاي كلاسيك از مجموعه‌هاي «نامعين»، آن‌ها را مجموعه‌هاي «معين»*[[11]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn11%22%20%5Co%20%22)* ناميده‌اند. مفهوم اخير را براي نخستين بار «لطفي‌زاده»*[[12]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn12%22%20%5Co%20%22)* در مقاله‌اي كه در سال 1965 منتشر ساخت بيان داشت. وي در اين مقاله تعريف و اهميت مجموعه‌هاي نامعين را اين‌گونه شرح داد:

معمولاً رده‌هاي اشياي موجود در جهان واقعي، معيارهاي دقيق و معين عضويت ندارند. مثلاً ردة حيوانات مشخصاً شامل سگ‌ها، اسب‌ها، پرندگان، و غيره به عنوان اعضاي آن مي‌شود و اشيايي چون سنگ‌ها، مايعات، گياهان و... بدان تعلق ندارند. با اين‌حال، هويت‌هايي چون ستاره‌هاي دريايي، باكتري‌ها و ... در ردة حيوانات، جايگاهي مبهم دارند. چنين ابهامي را براي عددي چون 10 در رابطه با «رده» تمامي اعداد حقيقي بسيار بزرگتر از1 مي‌توان ديد.

مسلماً «گروه تمامي اعداد بسيار بزرگتر از يك» يا «ردة زنان زيبا» و يا «ردة مردان بلندقد» از نقطه‌نظر اصطلاحات معمول رياضي، مجموعه‌ يا رده‌ به حساب نمي‌آيند؛ اما حقيقت آن است كه چنين «رده‌هايي» با تعريفي مبهم و نامشخص در انديشة بشري، بويژه در حوزه‌هاي شناخت الگو، ارتباطات اطلاعات و انتزاع نقشي مهم دارند.

«لطفي‌زاده» مجموعة نامعين را «رده‌اي از اشيا با درجات عضويت پيوستاري» مي‌داند و اظهار مي‌دارد كه «يك چنين مجموعه‌اي با يك عامل عضويت (تخصيص) كه درجه‌اي از عضويت بين صفر و يك را به هر چيزي اختصاص مي‌دهد، شناخته مي‌شود» (ص. 338). يك مجموعة A (مثلاً ردة مردان بلندقد در نمونه‌هاي «لطفي‌زاده») در نظرية مجموعه‌ها زيرگروهي از مجموعة مرجع جهاني X (مثلاً ردة تمامي مردان) است. براساس نظر «لطفي‌زاده» به هر عضو x از مجموعة جهاني X با استفاده از عامل عضويت، «درجة عضويتي» در مجموعة A اختصاص مي‌يابد كه از صفر (نه A) تا يك (A) در نوسان است. او خاطرنشان مي‌سازد كه مجموعه‌هاي نامعين موردنظر وي، متفاوت از نوع متداول مجموعه‌هاي معين است كه در آن‌ها، تابع عضويت صرفاً برحسب تعلق يا عدم تعلق x به مجموعة A (كه با دو ارزش يك و صفر است) شناخته مي‌شود.

در عملگرهاي مجموعة كلاسيك با دو مجموعه، فرض بر آن است كه مورد مشاهده بطور مجزايي به هر دو مجموعه تعلق دارد؛ مثلاً، انسان مي‌تواند همزمان هم دختر و هم دانشجو باشد و بر اين اساس، مي‌تواند در هر دو مجموعة همپوشان فوق جاي داشته باشد. ولي در مجموعه‌هاي نامعين، عضويت ناواضح است و اين مي‌تواند تمام انواع متغيرهاي بيروني را به صحنه آورد. با توجه به مثال بالا، [از نظر تئوري مجموعه‌هاي نامعين]، دختر موردنظر درواقع دو جنسي است. از همين يك مورد مي‌توان دريافت كه چگونه مجموعه‌هاي نامعين، روابط را پيچيده مي‌كند.

«بايس» و «ميدو» و «كرافت»*[[13]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn13%22%20%5Co%20%22)* در كتاب خود درباب سنجش در علم اطلاع‌رساني مي‌نويسند: «مسئلة اساسي در اندازه‌گيري مجموعه‌هاي نامعين، ارزشگذاري ميزان عضويت در مجموعه است» از نظر اين نويسندگان، چنين امري كم‌وبيش اختياري است. آنان براي نمايه‌سازي مدارك دو روش اساسي را توصيه مي‌كنند: نمايه‌سازي ذهني توسط نمايه‌سازان و نمايه‌سازي عيني و عملي توسط نرم‌افزارهاي رايانه‌اي بر اساس فراواني كلمات در متن.

خود «برادفورد» در گزارش قانون پراكندگي، براي به‌دست آوردن تابع عضويت متناسب با اين قانون، از نمودهاي عيني استفاده كرد. اين نمودها در شيوة به‌كار رفته براي تعيين رده‌هاي a و b وc در جدول 1، آشكارا موجودند. با بررسي اين روند، مي‌توان به شيوة وي پي‌برد كه عبارت است از تقسيم كل ارجاعات هر مجله به تعداد سال‌هاي دربرگيرندة نمونه (چهار سال براي ژئوفيزيك كاربردي) به منظور تعيين تعداد كل ارجاعات در هر سال. بر اين اساس، رده‌هاي معين‌شده به صورت زيرقابل تبيين‌اند:
a) مجلاتي كه سالانه بيش از 4 ارجاع دارند، b) مجلاتي كه سالانه حداقل 4 و حداكثر بيش از 1 ارجاع دارند، و c) مجلاتي كه سالانه 1 ارجاع يا كمتر دارند. پياده‌سازي اين روش به خارج قسمت‌هايي با حداكثر دو مرتبة اعشاري مي‌انجامد كه عامل ايجاد رده‌هايي با مرزبندي‌هاي زير است: بين a و b در 01/4؛ بين b و c در 01/1، و بين c و ردة صفر ـ كه من خود آن را اضافه كرده و d ناميده‌ام ـ در 01/0. اگر ردة a را كه شامل 8/2 درصد مجلات و 2/32 درصد ارجاعات در موضوع ژئوفيزك كاربردي است، در نظر بگيريم و آن را «هستة نشريات ادواري و بسيار مختص و مرتبط با موضوع» بدانيم، براي مجموعه‌هاي برادفورد، تابع عضويت زير قابل تخصيص است:

اگر تعداد ارجاعات سالانه به يك مجله از 4 بزرگتر باشد، نمرة/ درجة عضويت آن مجله برابر يك است. ولي اگر تعداد ارجاعات سالانه به يك مجله برابر 4 يا كمتر از آن باشد، درجة عضويت آن مجله برابر با تعداد ارجاعات به آن در سال، تقسيم بر 01/4 است.

عدد 01/4 با نشان‌دادن پايين‌ترين حد هسته به عنوان خارج قسمت در بخش دوم تابع عضويت انتخاب شده است. به‌كارگرفتن چنين تابع عضويتي براي داده‌هاي برادفورد، منتج به نتايج جدول 2 در موضوع ژئوفيزيك كاربردي مي‌شود. در اين جدول، رده‌هاي برادفورد به همراه ردة اضافه شدة صفر (d) بر اساس اصول مجموعه‌هاي نامعين كه به ترتيب، تنزل عضويت مجموعه را مي‌نماياند، نامگذاري و نشان داده شده‌اند. A=a؛ A=b و نه A؛ c= نه A وA  و بالاخره =d نه A . بررسي اين جدول نشان مي‌دهد كه در پايين مجلات هسته يا همان ردة A، نمرة عضويت مجلات سريعاً كم مي‌شود، به طوري كه تعداد اين مجلات همزمان زياد مي‌شود و درنهايت، حجم عظيمي از مجلات، در واقع صرفاً به صورت حاشية مجموعة ژئوفيزك كاربردي به حساب مي‌آيند. تعداد مجلات موجود در ردة صفر (d) تعمداً مشخص نشده است؛ چه، اين خود سؤال پيچيده‌اي است كه «برادفورد» هم نتوانست با موفقيت به آن پاسخ گويد.

**استنتاج‌هاي آماري «پيرسون»*****[[14]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn14%22%20%5Co%20%22)***

وجود خصيصة نامعين‌بودن مجموعه‌هاي برادفورد، در كاربرد روش‌هاي آماري براي تحليل كتابخانه تاثير بسزايي دارد. اين تاثيرات برگرفته از اصل وحدت علمي است، همان چيزي كه «برادفورد» قانون خود را براساس آن بنانهاد. «برادفورد» اظهار داشت: «برطبق اين اصل، هر موضوع علمي، كم‌وبيش با ديگر موضوعات علمي پيوستگي دارد». برمبناي اين اصل، هر اندازه نمرة عضويت مدارك يا مجلات در يك مجموعة معين برادفورد كم شود، راه براي ديگر موضوعات علمي هموار مي‌شود و بدين‌شكل عوامل بيروني به اين مجموعه نفوذ مي‌كنند. اين امر منتج به ناهمگني ـ و چه بسا ناهمنگي شديد ـ در مجموعه‌هاي برادفورد مي‌شود.

**جدول 2. قانون برادفورد از ديد تئوري مجموعه‌هاي نامعين در موضوع ژئوفيزيك كاربردي، سال‌هاي 1931ـ1928**

|  |  |  |  |
| --- | --- | --- | --- |
| **رده‌ها** | **تعداد ارجاعات در سال** | **مجلات داراي ارجاعات** | **نمره عضويت** |
| ژئوفيزيك كاربردي (**a**) | 25/23 | 1 | 000/1 |
| 50/21 | 1 | 000/1 |
| 00/14 | 1 | 000/1 |
| 00/12 | 1 | 000/1 |
| 50/11 | 1 | 000/1 |
| 75/8 | 1 | 000/1 |
| 00/7 | 1 | 000/1 |
| 00/5 | 1 | 000/1 |
| 25/4 | 1 | 000/1 |
| مرز رده‌هاي **a** و **b** | 01/4 |  | 000/1 |
| ژئوفيزيك كاربردي و نه ژئوفيزيك كاربردي (**b**) | 00/4 | 4 | 988/0 |
| 75/3 | 1 | 935/0 |
| 50/3 | 5 | 873/0 |
| 00/3 | 1 | 748/0 |
| 75/2 | 2 | 686/0 |
| 50/2 | 5 | 62/0 |
| 25/2 | 3 | 591/0 |
| 00/2 | 8 | 499/0 |
| 75/1 | 7 | 436/0 |
| 50/1 | 11 | 374/0 |
| 25/1 | 12 | 312/0 |
| مرز رده‌هاي **b** و **c** | 01/1 |  | 252/0 |
| نه ـ ژئوفيزيك كاربردي و ژئوفيزيك كاربردي (**c**) | 00/1 | 17 | 249/0 |
| 75/0 | 23 | 187/0 |
| 50/0 | 49 | 125/0 |
| 25/0 | 169 | 062/0 |
| مرز رده‌هاي **c** و **d** | 01/0 |  | 002/0 |
| نه ـ ژئوفيزيك كاربردي (**d**) | 00/0 | ؟ | 000/0 |

اين ناهمگني آنچه را كه «كارل پيرسون» مسئلة بنيادي علم» خواند، با پيچيدگي مواجه مي‌سازد. «پيرسون» اساس اين مسئله را در ويراست سال 1911 دستور زبان علم، در مبحث عمومي مربوط به مفاهيم نوين احتمال و همبستگي مطرح ساخت و در رويكرد علمي خود با اشاره به اصل رده‌بندي نوشت: «رده‌بندي اشيا و شكل‌گيري قضاوت‌هاي محض برمبناي اين رده‌بندي ـ قضاوت‌هاي مستقل از انديشه‌هاي ذهني افرادـ اساساً هدف و شيوة علم نوين است». به زعم وي، تغيير، خصيصة اصلي واقعيت است و در ايجاد شرايط براي كاركردهاي علمي نقشي شايان توجه دارد. وي مي‌نويسد:

... اساس نتايج فيزيكدان يا شيميدان، تجارب متوسط‌اند كه دقيقاً منطبق بر هم نيستند. روية اداركي آن‌ها، حتي در مهم‌ترين شرايط هم تفاوت و تغيير معيني دارا است. هر چند آنان ممكن است اين تغييرپذيري را ناشي از خطاهاي مشاهده، ناخالص‌بودن نمونه‌ها، عوامل فيزيكي حاكم بر محيط و شرايط قلمداد كنند، اما اين تغيير وجود دارد و هرگاه اين تغييرپذيري را با رويه‌هايي مثل معدل‌گيري منتفي نماييم، از عرصة شهود به عرصة ذهن قدم مي‌گذاريم و به جاي جهان واقعي، جهاني ساختگي بنامي‌كنيم.

«پيرسون» بر اساس تغييرپذيري پديده‌ها، تئوري جديدي را طرح كرد كه برطبق آن، مقولة پيوستگي و ارتباط را جايگزين ايدة مرسوم علي‍ّت ساخت. وي نظرية جديد خود را با عبارات زير بيان داشت:

اگر به تشخيص فردي‍ّت در اركان كل هستي دست يابيم و همساني را وابسته به حدود و ظرافت طبقه‌بندي بدانيم، درمي‌يابيم كه علت و معلول... صرفاً حامل درجه‌اي از بار معنايي تشابه‌اند و نه تكرار مطلق. قانون علي‍ّت ساخته و پرداختة ذهن و مستخرج از پديده‌ها است و بيانگر ماهيت واقعي آن‌ها نيست. مسئلة اساسي فراروي بشر، چيزي بيش از «علي‍ّت» است و مي‌توان آن را چنين خلاصه كرد: اگر «علت‌ها» مشابهتي با درجات مختلف دارند، «معلول‌ها» تا چه حد همساني خواهند داشت؟ از ديدگاهي وسيعتر، چيزي را مي‌توان عل‍ّت يك پديده دانست كه پيش از آن يا همراهش مي‌آيد. اگر ما اين عل‍ّت را تغيير دهيم، تا چه حد به تغيير و تبديل پديده منجر مي‌شود؟ اگر اين تغيير، اثري بر آن پديده نداشته، در اين صورت استقلال مطلق حاكم است. اگر چنين برداشت كنيم كه صرفاً تغييرپذيري عل‍ّت، عامل دگرگوني پديده بوده‌، در اين صورت وابستگي مطلق حاكم است. اين‌گونه از وابستگي مطلق يك پديده به يك عل‍ّت مشخص و سنجش‌پذير، از استثناها است... چنين تعامل و ارتباطي صرفاً در محدودة ذهني مصداق دارد و در عينيت آن در جهان واقعي جداً شك داريم؛ اما نمود تمامي درجات ارتباط و پيوستگي بين دو محدوده ـ يعني استقلال مطلق و وابستگي مطلق ـ امكان‌پذير است.

«پيرسون» با مطرح‌كردن «مسئلة بنيادي علم» در حيطة اين نوع رده‌بندي و تغييرپذيري چنين نوشت: جهان از بيشمار شيء تشكيل شده كه هويت مجزا و غيرپاياي فرضي دارند. همه مي‌توانند با اندازه‌گيري و مشاهدة خصيصه‌هاي اين اشيا آن‌ها را در رده‌هايي با هويت‌هاي مشابه جاي دهند. تغييرپذيري را در اين رده‌ها مي‌توان مورد توجه قرار داد و مسئلة اساسي علم، يافتن چگونگي همبستگي و وابستگي در تغيير يك رده به ردة ديگر است.

«پيرسون» اين نظرية جديد را با ترسيم متغير A در مقابل متغير B در يك نمودار پراكندگي نشان داد. نقاط روي نمودار به شكل عمومي يك منحني پراكنده شده‌اند. برطبق نظر وي، با عسكبرداري نمودار فوق از فاصلة 50 ياردي يا با مشاهدة آن از طريق تلسكوپ وارونه، يك فيزيكدان مي‌تواند به بررسي آن بپردازد. با كاربرد چنين روش‌هايي، نقاط پراكنش به صورت منحني منظم درمي‌آيند و بدين‌گونه، مشاهدة واقعي، جايگزين تابع رياضي مي‌شود. «پيرسون» ارتباط علي‍ّت با همبستگي را با استفاده از نمودار پراكنش چنين خلاصه كرد:

دو ردة قابل سنجش از اشيا را در حيطة فيزيك، شيمي، اجتماعي يا اقتصاد به صورت ادراكي در نظر بگيريد. در اين موارد هم، اين نقاط و نمودار پراكندگي وجود دارد. در بعضي موارد، اين نقاط در سراسر صفحه پراكنده مي‌شوند كه در اين صورت، همبستگي بين A و B وجود ندارد. در مواردي هم شكل نواري گسترده دارند، كه نشانگر ارتباطي متوسط است. وقتي نقاط به صورت ستارة دنباله‌دار درمي‌آيند، ارتباط بيشتر مي‌شود. با اين حال، تمام رديف‌هاي نمودارها، پيوستاري است كه نمي‌توان خطي به نشانة پايان همبستگي و شروع علي‍ّت در جايي از آن‌ها مشخص كرد. علي‍ّت صرفاً محدودة ذهني همبستگي است كه در حين باريك‌شدن بيش از حد نوار نمودار ظاهر مي‌شود و شبيه منحني به نظر مي‌رسد.

«پيرسون» با انعكاس ديدگاه‌هاي خود در نظرية جديدش، رويكرد قديمي عل‍ّت و معلول را اين‌گونه نكوهش كرد:

... هر نوع تغييري در اجزاي موجود يك رده با تغييري متناظر در بين اجزاي ردة ديگري مرتبط است. علم بايد بتواند ميزان شدت و ضعف اين تغييرات همگام را اندازه‌گيري كند. در اين محدودة ارتباط ذهني، استقلال مطلق در يك انتها و وابستگي مطلق در انتهاي ديگر آن قرار دارند. ديدگاه قديمي علت و معلول در پي گنجاندن جهان تحت اين دو محدودة ذهني است. چنين چالشي محكوم به شكست است. در مشاهدات عيني ما، اشيا يا مستقل يا عل‍ّي نمي‌توانند باشند.

با توجه به اينكه «پيرسون» هم در ضريب احتمال خود و هم در نسبت همبستگي خود به استقلال مطلق، عدد صفر و به وابستگي مطلق، عدد يك را تخصيص داده است، مسلماً به نامعين‌بودن محدودة بين صفر و يك، يعني ناحية نامعين ميانگين خروج از عضويت توجه داشته است.

**«داده‌هاي پرت» از ديد «بارنت» و «لوئيس»*****[[15]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn15%22%20%5Co%20%22)***

نامعين‌بودن مجموعه‌هاي برادفورد به خاطر وجود سازوكار داده‌هاي پرت، *[[16]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn16%22%20%5Co%20%22)* روند استنباط آماري «پيرسون» را از طريق اندازه‌گيري تأثير يك مجموعه بر ديگري پيچيده كرده است. «بك‌من»*[[17]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn17%22%20%5Co%20%22)* و «كوك»*[[18]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn18%22%20%5Co%20%22)* داده‌هاي پرت را يك مفهوم ذهني و پسداده‌‌اي مي‌دانند و «بارنت» و «لوئيس» نيز در كتابي آن را تاييد كرده‌اند. آنان در اين كتاب، داده‌هاي پرت را در مجموعه‌اي از داده‌ها چنين تعريف كرده‌اند: «مشاهده‌اي (يا زيرمجموعه‌اي از مشاهداتي) كه با بقية آن مجموعه از داده‌ها ناسازگار باشد». سپس آنان موضوع بحراني مربوط به داده‌هاي پرت را اين‌گونه پيش كشيدند:

عبارت «ناسازگار باشد» مهم و قابل توجه مي‌باشد و مقولة ذهني مربوط به مشاهده‌كننده است، چه يك مشاهده (يا مجموعه‌اي از مشاهدات) براي وارسي دقيق مشاهده مشخص شده باشند و چه نشده باشند. اما اين‌كه آيا بعضي مشاهدات، اعضاي واقعي جمعيت اصلي‌اند، يا آلوده‌كننده*[[19]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn19%22%20%5Co%20%22)* (يعني برخاسته از بعضي توزيع‌هاي ديگر) هستند، مهم مي‌باشد. مورد اخير ممكن است استنتاج‌هاي برگرفته از جمعيت اصلي را بي‌نتيجه سازد.

«بارنت» و «لوئيس» مسئلة داده‌هاي پرت را به فرضيه‌هاي احتمال توزيع واقعي جمعيت، كاملاً مرتبط دانستند. مشاهده‌اي كه ممكن است بر طبق فرضية توزيع نرمال جزو داده‌هاي پرت باشد، در صورتي كه مشاهده‌گر انتظار تويع كاملاً منحني‌وار در داده‌هاي زيستي، اجتماعي و اطلاعاتي را داشته باشد، مسئلة خاصي پيش نمي‌آورد. بنابراين آنان فرضيه‌هاي صفر*[[20]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn20%22%20%5Co%20%22)* و كاري*[[21]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn21%22%20%5Co%20%22)* هر آزمون ناسازگاري براي داده‌هاي پرت را نوعي مدل اساسي احتمال براي ايجاد تمام داده‌ها، بدون توجه به داده‌هاي پرت مي‌دانند. اگر شواهدي مهم براي رد فرضية كاري پيدا شود، «بارنت» و «لوئيس» الگوهاي «آلودگي» يا الگوهاي ناشي از «داده‌هاي پرت»*[[22]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn22%22%20%5Co%20%22)* را تبيين مي‌كنند كه ممكن است به عنوان فرضيه‌هاي جانشين وارد عمل شوند.

دو تا از اين فرضيه‌ها بيشترين ارتباط را با اين مقاله دارند. به اولين نوع اين فرضيه‌ها نام «جانشين جبرگرايانه»*[[23]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn23%22%20%5Co%20%22)* دادند كه شامل مواردي از داده‌هاي پرت مي‌باشد كه از اشتباهات فاحش انساني در اندازه‌گيري، ثبت و ضبط و ... ناشي شده‌اند. دومين واژة «بارنت» و «لوئيس» «جانشين تركيبي»*[[24]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn24%22%20%5Co%20%22)* نام دارد و بر آن است كه نمونة تحت بررسي، داراي آلودگي‌هايي جدا از مدل اساسي جمعيت است و اين اعضاي «خارجي» نمونه يا همان آلاينده‌ها خود را در قالب داده‌هاي پرت مي‌نمايانند. با تنزل سريع درجة عضويت در اعضاي مجموعة برادفورد كه به سرعت به سوي چنين آلودگي‌هايي پيش مي‌روند، جانشين تركيبي اهميت بيشتري در تحليل آماري داده‌هاي كتابخانه پيدا مي‌كند.

**تصوير عملي**

اكنون در قالب داده‌هاي منتج از پروژة تغيير ساختار موجودي پيايندها در دانشگاه ايالتي لويزيانا، *[[25]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn25%22%20%5Co%20%22)* تصويري از مفاهيم فوق را ارائه مي‌كنم. در بخشي از مقدمات اين پروژه، نياز هيئت علمي بخش شيمي دانشگاه لويزيانا در آوريل 1993 به پيايندها مورد بررسي قرار گرفت. لازم به ذكر است كه فقط نياز هيئت علمي بخش شيمي مورد بررسي قرار گرفت و بخش‌هاي بيوشيمي و مهندسي شيمي بررسي نشدند. از هيئت علمي اين دانشگاه خواسته شد بدون اين‌كه خود را محدود به پيايندهاي مورد اشتراك دانشگاه كنند، پيايندهاي مهم در امر تحقيق و تدريس خود را مشخص سازند. نشريات منتخب آنان بر اساس موضوعات مندرج در Science Citation Index Journal Citation Reports (SCIJCR)، در سال 1993 تقسيم‌بندي گرديد (اين نمايه‌نامه را «مؤسسه اطلاعات علمي»*[[26]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn26%22%20%5Co%20%22)* منتشر مي‌كند).

پيايندهايي كه هيئت علمي بخش شيمي دانشگاه لويزيانا انتخاب كرده بودند و همراه با قانون پراكندگي برادفورد برطبق موضوعات متعدد (موجود در مؤسسة اطلاعات علمي) مرتب شده بودند، از جمله اين موارد را شامل بودند: مهندسي برق و الكترونيك، علم محيط، علم زمين‌شناسي، علم مواد (سراميك)، تغذيه و رژيم غذايي، فيزيك، راديولوژي و پزشكي هسته‌اي.

با توجه به اين امر، لازم بود نمونه را به پيايندهايي كه «مؤسسة اطلاعات علمي» آن‌ها را در شاخه‌هاي مختلف شيمي طبقه‌بندي كرده بود (شامل مهندسي شيمي و كريستالوگرافي) محدود ساخت. استثنائاً موضوع اسپكتروسكوپي نيز به خاطر تأكيد بخش شيمي دانشگاه، به اين فهرست افزوده شدـ هر چند اين رشته عمدتاً بخشي از مبحث نور در فيزيك است. نتيجة نهايي نمونه‌اي شامل 154 مجله بود.

براي تعيين ارزش علمي اين 154 مجله، سه متغي‍ّر كم‍ّي به كار گرفته شد: نمرات هيئت علمي دانشگاه لويزيانا، استنادهاي كلي «نمايه‌نامة استنادي علوم» در سال 1993 و ضريب تأثير اين نمايه‌نامه در سال 1993. از اين مقادير فقط دو مورد اول معتبر شناخته شدند. امتيازات علمي دانشگاه لويزيانا به عنوان معيار اصلي اندازه‌گيري ارزش علمي به كار رفت، چرا كه منطق مجموعة مجلات با ارزيابي بخش شيمي و نيز با دلايل منطقي معين شده بود. در اين روش از مدرسين شيمي خواسته شد 10 عنوان مجله را نام ببرند و بيان كنند كه آيا لازم است اين مجلات مورد اشتراك دانشگاه باشند يا از طريق تحويل مدارك از راه دور در دسترس قرار گيرند؛ سپس عناوين مجلات را به ترتيب نزولي از 10تا 1 بر اساس رتبة داده‌شده به آن‌ها مرتب كنند. يك عنوان بدين‌ترتيب رتبه‌بندي شد: 10 نمره به هر بار انتخاب يك عنوان، 10 نمره به اين‌كه آن مجله در آن دانشگاه وجود داشته باشد، و از 10 تا يك نمره بسته به رتبه‌اي كه استاد به آن مجله داده بود. 25 استاد شيمي به اين تحقيق پاسخ دادند و 154 مجله در نمره‌دهي هيئت علمي از 10 تا 755 امتياز گرفتند (كه 755 متعلق به Journal of American Chemical Society بود). براي ارزيابي رتبه‌دهي هيئت علمي، آن‌ها را با استنادهاي كلي «نمايه‌نامة استنادي علوم» تطبيق دادند تا ميزان همخواني نمره‌دهي اساتيد دانشگاه لويزيانا با نظرات بخش انتشارات جامعة علمي مشخص شود. در حين انجام اين كار، با مشكل جدي «داده‌هاي پرت» روبرو شدم. جدول 3، پراكندگي دو متغير پيشگفته را در دوره‌هاي مشخص‌شده بر حسب چارك نشان مي‌دهد. برطبق اين جدول با نوعي توزيع برادفورد سروكار داريم كه در آن، انتهاي فوقاني توزيع بيشترين ارزش را دارد. بنابراين در مجموعه ژئوفيزيك كاربردي برادفورد، 8/2 درصد از مجلات مختص 2/32 درصد از مراجع هستند ودر اينجا، ردة مربوط به چارك فوقاني، 5/62 درصد از 154 عنوان مجله را شامل مي‌شود كه تركيبي از نمره‌دهي هيئت علمي و 2/80 درصد از استنادهاي كلي آن‌ها است.

**جدول 3. توزيع 154 عنوان مجله شيمي با ترتيب نزولي برطبق نمره‌دهي هيئت علمي و استناد كلي «نمايه‌نامة استنادي علوم» در رديف‌هاي مشخص‌شده برحسب چارك**

|  |  |  |
| --- | --- | --- |
|  | **نمرات هيئت علمي** | **اسنادهاي كلي «نماية استنادي علوم»** |
| **محدوده گروه چارك** | **درصدنمرات هيئت علمي** | **محدودة گروه چارك** | **درصد استنادهاي كلي «نماية استنادي علوم»براي مجلات** |
| **بالا** | 755ـ111 | 5/62 | 231324ـ11685 | 2/80 |
| **بالاي مياني** | 110ـ50 | 6/20 | 11586ـ3303 | 4/13 |
| **پايين مياني** | 50ـ33 | 1/11 | 3285ـ1533 | 6/4 |
| **پايين** | 32ـ10 | 8/5 | 1526ـ255 | 8/1 |

توزيع فراواني نمرات هيئت علمي و استنادهاي كلي، به صورت ترسيمي در نمودار 1 و 2 نمايش داده شده و به وضوح، شاهدي بر حضور آلايندة جمعيتي غير از آنچه براساس امتيازات اعضاي هيئت علمي ترسيم شده است را نشان مي‌دهد. آلاينده‌ها كه يك پديدة طبيعي در توزيع برادفورد هستند، با مشاهدة حد انتهاي راست مشخص نيستند، اما در موقعيت‌هاي نسبي مشابه مثل Journal of the American Chemical Society (با نمرة هيئت علمي برابر 755 و استناد كلي «نمايه‌نامة استنادي علوم» برابر 148900) و Journal of Biological Chemistry (با نمرة هيئت علمي برابر 197 و استناد كلي «نمايه‌نامة استنادي علوم» برابر 231324) چنين نيست. در نمودار 1 كه فراواني توزيع 154 مجله برحسب رتبه‌بندي هيئت علمي دانشگاه را نشان مي‌دهد، موقعيت مجله Journal of the American Chemical Society در سمت راست با منطق مجموعه همسان است. هرچند در نمودار 2 نشان‌دهندة توزيع فراواني اين مجلات برحسب استناد كلي «نمايه‌نامة استنادي علوم» است، مجله Journal of Biological Chemistry در سمت راست نشريه Journal of the American Chemical Society قرار گرفته و با منطق مجموعه مطابقت ندارد.      فراواني

Journal of
Biological Chemistry

Journal of
 AmericanChemical Society

نمرات اعضاي هيئت علمي

**نمودار 1. توزيع فراواني 154 عنوان مجله در موضوع شيمي بر حسب نمره‌دهي اعضاي هيئت علمي دانشگاه ايالتي لويزيانا**

فراواني

Journal of
 AmericanChemical Society

Journal of
Biological Chemistry

استناد كلي

**نمودار 2. توزيع فراواني 154 عنوان مجله در موضوع شيمي برحسب استنادهاي كل‍ّي «نمايه‌نامة استنادي علوم»**

مجله Journal of Biological Chemistry جزو «داده‌هاي پرت» محسوب مي‌شود. اين پيش‌بيني در نمودار 3 مورد تأييد قرار گرفته است كه توزيع پراكندگي نمرات هيئت علمي را در برابر استنادهاي كلي «نمايه‌نامة استنادي علوم» نشان مي‌دهد. در اين شكل يك خط رگرسيون فرضي از مجله Journal of the American Chemical Society تا مبدأ ترسيم شده كه از نقاط مياني مي‌گذرد، در حالي كه خط ترسيم‌شده از مجله Journal of Biological Chemistry تا مبدأ در قسمت پايين و سمت راست تمام نقاط قرار دارد. وقتي كه اين مجموعه را ساختم، تصميم گرفتم كه بيوشيمي را زيرشاخه‌اي از شيمي قلمدادكنم. تصميم من در اين زمينه از طرح رده‌بندي كتابخانة كنگره تأثير پذيرفته بود كه در آن، بيوشيمي به عنوان زيرمجموعه‌اي از شيمي آلي در شيمي است. با وجود اين، موقعيت مجله Journal of Biological Society مرا به شك واداشت. پژوهش بعدي اين سوءظن را تأييد كرد. برخلاف طرح رده‌بندي كنگره، در رده‌بندي دهدهي ديوئي، بيوشمي نه تنها شاخه‌اي از شيمي نيست، بلكه به عنوان زيرشاخه زيست‌شناسي و علوم زيستي تقسيم‌بندي شده است. نه تنها هيئت علمي دانشگاه لويزيانا بخش‌هاي شيمي و بيوشيمي را از هم جدا كرده‌اند، حتي در آخرين رتبه‌بندي برنامه‌هاي تحقيقاتي دكترا در انجمن ملي تحقيقات آمريكا*[[27]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn27%22%20%5Co%20%22)* نيز، شيمي زير سرفصل علوم فيزيكي و رياضيات رده‌بندي شده است و بيوشيمي ضمن تركيب با زيست‌شناسي مولكولي، زير سرفصل علوم زيستي رده‌بندي مي‌شود. بنابراين، Journal of Biological Society همراه با تعدادي از مجلات بيوشيمي ديگر در مجموعة موردنظر من، نتيجة نامعين مجموعه‌هاي برادفورد است. نشريه فوق هم A و هم نه A، هم شيمي و هم بيوشيمي است.

Journal of
 AmericanChemical Society

Journal of
Biological Chemistry

استناد كلي

**نمودار 3. رابطة بين رتبه‌بندي مجلات از ديد هيئت علمي دانشگاه لويزيانا و استنادهاي كلي «نمايه‌نامة استنادي علوم»**

**روش‌هاي مهاركردن داده‌هاي پرت**

«بارنت» و «لوئيس» روش‌هاي مهاركردن داده‌هاي پرت را در چهار گروه كلي بيان كرده‌اند: قوانين محكم و صريحي مبني بر اين‌كه از كدام گروه از روش‌ها استفاده ‌شود، وجود ندارد؛ بلكه اين امر به چگونگي پديدآمدن داده‌هاي پرت و هدف فرد بستگي دارد. «بارنت» و «لوئيس» يكي از مقوله‌هايشان را «رد»*[[28]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn28%22%20%5Co%20%22)* ناميدند. منظور آن‌ دو، عبارت بود از اين‌كه اگر داده‌هاي پرت قابل تصحيح نباشد، فرد آن‌ها را از دور خارج مي‌كند و سپس باقي نمونه را تحليل مي‌نمايد. هنگامي كه من پنج مورد از داده‌هاي پرت را در حين تعيين همبستگي توليد ـ زمان «پيرسون» بين رتبه‌بندي هيئت علمي و استنادهاي كلي «نمايه‌نامة استنادي علوم» پيدا كردم، چنين موردي پيش آمد. تحليل بقية موارد، پنج دادة پرت را (كه از بين آن‌ها چهار مورد از نظر هيئت علمي نسبت به استنادهاي كلي، رتبة پايين داشتند) ايجاد نمود كه دو مورد از داده‌هاي پرت اخير را «مؤسسة اطلاعات علمي» در بيوشيمي مولكولي رده‌بندي كرده است. ضريب همبستگي اوليه 66 درصد بود كه پس از حذف داده‌هاي پرت از نمونه، اين ضريب به 72 درصد رسيد.

دومين گروه در روش‌هاي مهاركردن داده‌هاي پرت برطبق نظر «بارنت» و «لوئيس»، «تعيين هويت»*[[29]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn29%22%20%5Co%20%22)* ناميده شده. منظور آن‌ها اين بود كه داده‌هاي پرت ناسازگار، علائمي از بعضي عوامل مفروض در حين كار بر روي جمعيت تحت تجزيه و تحليل هستند. من هم اين كار را انجام دادم و به نتايج جديدي در تبيين رابطة بيوشيمي با شيمي رسيدم. هرچند ابتدا فكر كرده‌ بودم بيوشيمي شاخه يا زيرمجموعه‌اي از شيمي است، اكنون آن را به عنوان مجموعه‌اي جدا با الگوهاي آماري مربوط به خود محسوب مي‌كنم.

گروه ديگر در روش‌هاي مهار داد‌ه‌هاي پرت از نظر «بارنت» و «لوئيس»، «الحاق»*[[30]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn30%22%20%5Co%20%22)* نام دارد. در اين روش، هدف آن است كه يك الگوي همگن را با الگوي همگن ديگر براي كل نمونه‌ها جايگزين كنيم (يكپارچه‌سازي داده‌هاي پرت) تا هيچ مشاهده‌اي داراي تناقض نباشد.

چهارمين و آخرين مقوله در روش‌هاي مهاركردن داده‌هاي پرت كه «بارنت» و «لوئيس» آن را پيشنهاد كردند، «سازگاري» نام دارد. آنان اين مقوله را به دو جزء تقسيم كردند: اولين جزء روش‌هاي باثبات هستند كه ارزش منطقي را در برابر داده‌هاي پرت حفظ مي‌كنند. يك نمونه از اين روش‌ها را مي‌توان براي آزمون استقلال مربع كاي*[[31]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftn31%22%20%5Co%20%22)* ـ به جاي استفاده از روش‌هاي همبستگي ـ براي تحقيق دربارة ارتباط نمره‌دهي هيئت علمي دانشگاه لويزيانا با استناد كلي «نمايه‌نامة استنادي علوم» به كار برد. آزمون استقلال مربع كاي را، «كارل پيرسون» اساساً برطبق احتمال پيشنهاد كرد. از آن‌جا كه روش‌هاي همبستگي مستلزم كاربرد رياضيات است كه هم تناسب نقاط داده‌ها را با خط رگرسيون در رابطة همبستگي ضريب گشتاوري پيرسون و هم رابطة يك رتبه با رتبه ديگر را در همبستگي رتبه‌بندي اسپيرمن، تعيين مي‌كند، بنابراين وابستگي متغيرها به يكديگر در گروه‌هاي بزرگ با استفاده از آزمون استقلال مربع كاي امكان‌پذير است. اين امكان در جدول 3 نشان داده شده است كه در آن، دو مجلة مذكور برحسب هر دو اندازة كيفيت علمي در چارك بالايي قرار گرفته‌اند.

جزء ديگر الحاق، حاوي آن روش‌هايي است كه با تخصيص ارزش كمتر به مقادير انتهايي در مقايسه با ديگر اعضاي نمونه، داده‌هاي پرت را مهار مي‌كنند. يك چنين روشي مي‌تواند winsovization باشد، كه در آن مشاهدات با نزديكترين مجاور آن جايگزين مي‌شود. در اين روش استنادهاي كلي «نمايه‌نامة استنادي علوم» از مجله Journal of Biological Chemistry به اندازة استنادهاي كلي همين نمايه‌نامه از مجله Journal of American Chemical Chemistry است.

با اين حال، «الحاق» شايد روش بهتري باشد؛ چرا كه نظرية مجموعه‌هاي نامعين را براي مهار داده‌هاي پرت به‌كار مي‌گيرد. كاربرد اين نظريه در طبيعت عملي است و به منطق مجموعة تحت بررسي و نيز هدف پژوهش بستگي دارد. هدف عمده از اين كار، سازگاركردن داده‌هاي پرت به تناسب عضويت آن‌ها در مجموعه است. در مثال فوق يكي از مصاديق نشان‌دهندة اين امر آن است كه نمره‌دهي هيئت علمي را به‌صورت تجربي، استخراج و از نتايج ميزان عضويت براي انطباق با استنادهاي كلي «نمايه‌نامة استنادي علوم» استفاده كنيم. يك روش ديگر، تجزيه و تحليل پايگاه «مؤسسة اطلاعات علمي» و محدودساختن استنادها به انواع متناسب با منطق مجموعه است.

بر اين اساس، نظرية احتمالات و نظرية مجموعه‌هاي نامعين متناقض نيستند، اما قالب‌هايي مكمل براي تجزيه و تحليل به حساب مي‌آيند.

**منابع و يادداشتها**

Bradford, S. C. "Sources of Information on Specific Subjects". **Engineering** 137:85-86 (1934).

A full treatment of the various formulations of Bradford’s Law is outside the scope of this paper.

Kosko, B. **Fuzzy Thinking: The new Science of fuzzy Logic**. London: Flamingo, 1994, p.6.

Bradford, S. C. "Some General Principles of a Bibliographical Classification Scheme, with Application to the Universal Decimal Classification". **Proceedings of the British Society for International Bibliography** 6 (3): 57-69 (1944).

Zadeh, L.A. "Fuzzy Sets". **Information and Control** 8 (3): 338-353 (1965).

Boyce, B.R., C.T. Meadow. And D.H. Kraft. **Measurement in Information Science**. San Diego: Academic Press, 1994, p. 95.

Bradford, S.C. (1948). "Complete Documentation" in **The Royal Society Empire Scientific Conference, June-July** 1946: Report. London: The Royal Society, 1948, Vol. 1, pp. 729-748.

Pearson, K. **The Grammar of Science**. 3rd ed., rev. and enl. London: A. and C. Black, 1911. Pt. 1, p. 165.

Beckman, R. J. and R.D. Cook. "Outliers". **Technometrics** 25 (2): 119-149 (1983).

Barnett, V. and T. Lewis. **Outliers in Statistical Data** 3rd ed. Chichester: J. Wiley, 1995. p. 7. A short, general review of outliers and the procedures for handling them is: Barnett. V. "The Study of Outliers: Purpose and Model". **Applied Statistics** 27 (3): 242-250 (1978)

Bensman, S.J. "The Structure of the Library Market for Scientific Journals: The Case of Chemistry". **Library Resources & Technical Services** 40 (2): 145-170 (1996): Bensman, SJ. And S.J. Wilder "Scientific and Technical Serials Optimization in an Inefficient Market: A LSU Serials Redesign Project Exercise". **Library Resources & Technical Services** 42 (3): 147-242.

**Science Citation Index Journal Citation Reports: A Bibliometric Analysis of Science Journals in the ISI Database**: 1993. Philadelphia: Institue for Scientific Information, 1994.

Glodberger, M.L., B.A. Maher, and P.E. Flattau, eds. **Research Doctorate Programs in the United States: Continsuity and Change**. Washington, D.C.: National Academy Press, 1995.

*[[1]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref1%22%20%5Co%20%22)*. Bensman, Stephen J. "Bradford’s law and fuzzy sets: statistical implementations for library analyses". IFLA Journal, 27(4). 2001. P: 283-46.

2. دانشجويان كارشناسي ارشد كتابداري و اطلاع‌رساني پزشكي

*[[3]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref3%22%20%5Co%20%22)*. Samuel Clement Bradford

*[[4]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref4%22%20%5Co%20%22)* . Kensington

*[[5]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref5%22%20%5Co%20%22)*. Science Museum Library

*[[6]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref6%22%20%5Co%20%22)*. fuzzy sets

*[[7]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref7%22%20%5Co%20%22)*. Kosko

*[[8]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref8%22%20%5Co%20%22)*. Aristotle

*[[9]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref9%22%20%5Co%20%22)*. British Society ofInternational Bibliography

*[[10]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref10%22%20%5Co%20%22)*. George Boole

*[[11]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref11%22%20%5Co%20%22)*. crisp

*[[12]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref12%22%20%5Co%20%22)*. Lotfi Zadeh

*[[13]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref13%22%20%5Co%20%22)*. Bayce, Meadow and Kraft

*[[14]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref14%22%20%5Co%20%22)*. Karl Pearson

*[[15]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref15%22%20%5Co%20%22)*. Barnet and Lewis

*[[16]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref16%22%20%5Co%20%22)*. outliers

*[[17]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref17%22%20%5Co%20%22)*. Beckman

*[[18]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref18%22%20%5Co%20%22)*. Cook

*[[19]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref19%22%20%5Co%20%22)*. contaminant

*[[20]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref20%22%20%5Co%20%22)*. null hypothesis

*[[21]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref21%22%20%5Co%20%22)*. working hypothesis

*[[22]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref22%22%20%5Co%20%22)*.outlier-generating

*[[23]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref23%22%20%5Co%20%22)*.uterministic alternative

*[[24]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref24%22%20%5Co%20%22)*. mixture alternative

*[[25]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref25%22%20%5Co%20%22)*. Louisiana State University (LSU)

*[[26]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref26%22%20%5Co%20%22)*. Institute for Scientific Information (ISI)

*[[27]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref27%22%20%5Co%20%22)*. US Research Doctorate Programs

*[[28]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref28%22%20%5Co%20%22)*. re-jection

*[[29]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref29%22%20%5Co%20%22)*. identification

*[[30]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref30%22%20%5Co%20%22)*. incorporation

*[[31]](http://www.aqlibrary.org/editor/main.htm%22%20%5Cl%20%22_ftnref31%22%20%5Co%20%22)*. chi-square