

# Kütüphanecilik ve Bilgibilim Alanlarında İstatistik Yöntemlerin Kullanımı\*

## The Use of Statistical Methods in Library and Information Science

Aydın Erar\*\*

### Öz

*Çalışmada, Kütüphanecilik ve Bilgibilim alanlarında yayınlanan dergilerde yer alan makalelerde hangi İstatistik yöntemlerin daha yoğun olarak kullanıldığı araştırılmıştır. Bu amaçla, SSCI (Social Science Citation Index)'de "Library Science and Information Science" kategorisinde yer alan dergilerden belirlenen makaleler üzerinden yapılan bir taramanın sonuçları sunulacaktır. Çalışma, 1991-97 arasında etki faktörleri (impact factor) en yüksek olan dergilerde yer alan toplam 1864 makale üzerinden yapılmıştır. İlk kısımda, bazı istatistik terimlerinin kısa bir özeti ve yöntemlerin yaklaşık bir sınıflaması verilmiştir. Yöntemlerin ana konulara göre dağılımları, her bir ana konudaki alt konuların sıralamaları sonucunda, makalelerin yaklaşık %69'unda nicel yöntemlere (Betimsel ve İleri İstatistik, Yöneylem Araştırması, Matematik) başvurulduğu; makalelerin %32'sinde ileri istatistik yöntemlerin kullanıldığı görülmektedir. Kütüphanecilik ve Bilgibilim alanlarında, "Çok Değişkenli İstatistik Yöntemler" ve "Stokastik Süreçler" konularının giderek artan bir yoğunluğu sahip olduğu gözlenmektedir.*

### Abstract

*The aim of this study is to call attention to the Statistical Methods used in Library and Information Sciences. Results of a search are given over journals,*

\* Bu çalışma, 10 Şubat 2000 tarihinde Isparta Süleyman Demirel Üniversitesi'nde yapılan AB2000 Akademik Bilişim Kongresi'nde sunulan bildirinin gözden geçirilmiş biçimidir.

\*\* Doç. Dr. Aydın Erar, Hacettepe Üniversitesi İstatistik Bölümü Öğretim Üyesidir, H.Ü. İstatistik Böl., e-posta: erar@hacettepe.edu.tr

found in "Library Science and Information Science" Category in SSCI, in order to reveal which Statistical Methods are used more frequently. This work was done using a total of 1864 papers that appeared in the highest impact factored journals between 1991-97. In the first part, a brief summary of some basic statistical terms and a subjective classification of methods is given. After the restrictions of this study are mentioned, distribution of Methods among Journals, distribution of Advanced Methods according to main topics of Statistics, an ordering of sub-topics in each heading are given. As a result, in approximately 69% of papers, quantitative methods (Descriptive and Advanced Statistics, Operation Research and Mathematics) have been used. 32% of all papers have used advanced statistical methods. Besides, it was observed that the subjects Multivariate Statistical Methods and Stochastic Processes have an increasing concentration in Library and Information Science.

### **Anahtar Kelimeler:**

*İleri İstatistiki Yöntemler, Nicel Yöntemler, Bilgibilim*

### **Keywords:**

*Advanced Statistical Methods, Quantitative Methods, Information Science*

## **Giriş**

Bu çalışmanın amacı, Kütüphanecilik ve Bilgibilim alanlarında kullanılan istatistik yöntemlere dikkat çekmektir. Burada, hangi istatistik yöntemlerin daha yoğunluklu kullanıldığını araştırmak üzere Kütüphanecilik ve Bilgibilim kategorisinde yer alan dergilerden yapılan bir tarama çalışmasının sonuçları verilecektir. Amaç, gerek araştırmalarda gerekse istatistik derslerinde alışlagelen istatistik yöntemlerin yanında hangi yöntemlere ayrıca yer verilmesi gerektiği ya da öncelikle hangilerinin verilebileceği konusunda, bu alanda çalışanlara yardımcı olmak, bir fikir verebilmektir.

İstatistik, araştırma verilerinin düzenlenmesi, özetlenmesi, örnekleme ile elde edilen sonuçların evrene/bütüne genellenmesi ya da hipotezler ile ilgili karar verme yöntemlerini kapsayan bilim dalıdır. Houser L.J. ve Lanorich G.T.(1978: 188), İstatistiği, "bilimsel araştırma dilidir" ifadesiyle tanımlamıştır. Wyllys R.E.(1978: 12-13), bir alanda İstatistik biliminin kullanılmasının, o alanın bilimsel olarak yönlendirilme derecesini gösterebileceğini belirtmiştir. Wallace D.P.(1985: 408), "bir araştırmanın bilimsellik düzeyi, çıkarılmalı (inference) istatistik yöntemleri kullanmaya olan eğilimi ile ölçülür" demektedir.

İstatistik, Blalock (1979: 1-10)'a göre, tanıtlanmak istenen bir şeyi tanıtlayabilen bir yöntem değildir; ancak, genelde, veri sınırları içinde açıklamalar yapacak yöntemlerdir. İstatistik, ölçme kurallarını koyan, görüşmecinin nasıl özenli olabileceğini belirten bir araç değildir. İstatistik, tek tek olayları ya da seyrek olarak ortaya çıkan olayları inceleyen bir dal da değildir. İstatistik, olayların tekrarlanabildiği, bu tekrarlarda ya da gözlemlerde değişkenliğin ortaya çıkabildiği yani istatistiksel verilerin elde edilebildiği durumlarda araştırmaya yön veren, sonuçlara güvenirlilik, inandırıcılık ve bilmisellik sağlayan yöntemler topluluğudur.

Pritchard(1969), yazılı iletişim ile ilgili aktivitelerin ve süreçlerin ölçülmesinde istatistiksel ve matematiksel yöntemlere gerek olduğunu belirtmiştir. Pao(1989: 5-7), bu yöntemlerin, mevcut bilgi sistemindeki eksikliklerin tanımlanması, bilgi hizmetlerinin daha iyi planlanması ve yönetimi, bilgi erişim sistemlerinin iyileştirilmesi, büyümenin görülmesi, ileriki kullanım ve eğilimlerin önceden kestirmesi için kullanılabilirliğini açıklamıştır. Hafner(1998: 1-50), kütüphane siyasasını biçimlendirmede, kısa/uzun dönemli stratejik planlama amaçlarını geliştirmede, kütüphanelerin performanslarını incelemeye, disiplinler arası geçiş eğilimlerinin görülmesinde, elit dergi ve/ya da yazarlara karar verilmesinde vb. bu yöntemlerden yararlanılabileceğini belirtmiştir.

Bu çalışmada bazı temel istatistik kavramlarının, bütünlük sağlamak amacıyla, kısa bir özeti verildikten sonra kapsam ve kısıtlar üzerinde durularak, sonuçlar gösterilecektir.

## Temel İstatistik Kavramları ve Bir Sınıflama

*Betimsel (Descriptive) İstatistik:* Verileri daha kullanışlı ve açıklayıcı duruma getirmek için bu verilerin düzenlenmesi ve özetlenmesidir. *Çıkarsamalı (Inferential) İstatistik,* bir bütünü (evreni) temsil edebilecek nitelikte örneklem seçerek, bu örneklem üzerinden elde edilen sonuçların bütüne genellenmesi ya da öne sürülen hipotezlerin doğruluklarının yine bu örneklem kullanılarak test edilmesi anlamında kullanılacaktır.

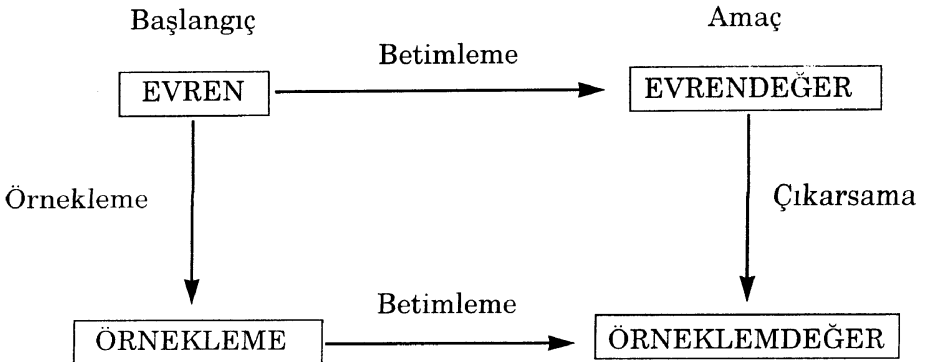
İstatistik'de araştırmaya konu olan en küçük birim *denek* ya da *gözlem*; tüm araştırma birimlerinin oluşturduğu topluluk ya da bütün *evren*; evreni temsil edebilecek nitelikte istatistik kurallara göre rasgele seçilen birimlerin topluluğu *örneklem*; evren üzerinden bulunan değerler *evrendeğer*; örneklem üzerinden bulunanlar da *örneklemdeğer* olarak adlandırılır.

İstatistik'de işlemler değişken diliyle yapılır. Değerleri birimden birime değişen her bir özellik *değişken* (variable) olarak tanımlanır. Yöntemlerin uygulaması değişken türlerine, yapılarına vb. göre değişir. Yaş, uzunluk gi-

bi değişkenler *nicel ve sürekli*, bir web sayfasını ziyaret eden birey sayısı *nicel ve kesikli*, her bireyin tercih ettiği arama motoru türü *nitel ve sınıfsal*, yazarların aldıkları atıf sayılarına göre sıralanışları *nitel ve sırasal* değişkenlerdir. Eğer bir değişken başka bir değişkenin etkisiyle değişiyor ise *bağımlı değişken*, başka değişkenlerin etkisinde değilse *bağımsız değişken* adını alır. Öte yandan her değişken değerine belli olasılıklar atanabiliyor ise bu değişkene *rastlantı değişkeni* denir. Örneğin değişken, okuyucuların aldıkları kitap sayısı olsun. Hiç kitap almama olasılığı, 0.10; 1 kitap alma olasılığı 0.60; 2 kitap alma olasılığı 0.20, vb. biçiminde tanımlar verilebiliyorsa, değişken rastlantı değişkenidir.

*Olasılık* kavramı, İstatistiğin temel konusu olduğu kadar bilgi sistemlerinde de sıkça kullanılır. Feeney ve Grieves (1994), bilgiyi, belirsizliği azaltan şey olarak tanımlarlar. Olasılık, belirsizliğin bir ölçüsüdür. Böylece iki kavram birlikte önemli bir anlam kazanır. Olasılık, bir olayın gerçekleşme şansının sayısal anlatımı olduğu kadar, evren üzerinden ya da uzun dönemde elde edilen sıklık (frekans) oranıdır. Belirsizliğin olasılık ölçüsünün elde edilmesi, olayların olasılıklarının hesaplanması, olasılık dağılımlarının araştırılmasını gerektirir. Burada *olasılık dağılımı*, evren üzerinden elde edilen sıklık dağılışıdır; kuramsal dağılımlardır. Bookstein (1988: 7)'e göre Bilgibilimi'nde birçok karar verme problemi aktiviteyi anlatan bir dağılıma bağlıdır. Bu dağılımların bilinmesi, olasılık hesaplarında, özetleme ölçülerinin hesaplanmasında, örneklemenin evreni temsil edip etmediğinin araştırılmasında, anormalliklerin saptanmasında, model kurmada, vb. yardımcı olur.

İstatistiksel verilerin değerlendirilmesinde başlangıç, evren ya da bütündür; amaç ise evrene ilişkin bilgilerin, bulguların yani evrendeğerlerin elde edilmesidir. Bu amaca ulaşmada izlenen bir yol, evrenin doğrudan ken-



Şekil 1. Bir araştırmadaki istatistiksel süreç

disini incelemektir. Eğer evren incelenemiyorsa ya da incelemek gereksiz ise örnekleme yoluyla seçilecek örnekleme incelemek ve evrene genelleme yapmaktır. Bu süreç aşağıda gösterilmektedir.

Şemada yer alan terimlerin genel kapsamları, bir sınıflama biçiminde, aşağıda verilmiştir.

Betimleme genel konuları:

- Tek Değişkenli Düzenleme ve Özetleme (Sıklık Çizelgeleri; Grafikler; Ortalamalar, Sapma Ölçüleri; Oranlar; vb.)
- İki Değişkenli Düzenleme ve Özetleme (Çapraz Çizelgeleri; Grafikler; Oranlar; İlişki; Doğrusal ve Doğrusal Olmayan Bağlıntılar; Zaman Dizileri vb.)
- Çok Değişkenli Düzenleme ve Özetleme (Çapraz Çizelgeler, Log-Doğrusal Modeller; Oranlar; İlişki; Bağlıntılar; Boyut Küçültme; Sınıflandırma; vb.)
- Olasılık, Olasılık Dağılımları ve Stokastik (Olasılıksal) Süreçler.

Örnekleme (Sampling) genel konuları:

- Alışlagelen örnekleme yöntemleri ve dağılımları
- Deneysel Tasarım (Experimental Design)
- Benzetim (Simulation)

İstatistiksel çıkarıma genel konuları (belli bir yanılma payı ya da risk altında):

- İstatistiksel Kestirim (Estimation) ile genelleme
  - \* Nokta Kestirimleri (En çok olabilirlik(EÇÖ), En küçük kareler(EKK), Bayes, Sağlam, vb.)
  - \* Aralık Kestirimleri
- Hipotez Testleri ve Karar Verme
  - \* Örnekleme değeri evrendeğeri temsil niteliğinin testi
  - \* Örneklem ya da gruplar arasındaki farklılıkların testleri
  - \* Örneklem dağılımlarının kuramsal dağılımlara uygunluklarının testleri
  - \* İlişki, birliktelik ya da bağıntılar üzerine testler
- Süreçler ile ilgili çıkarımlar

Öte yandan *Betimsel ya da Çıkarsamalı İstatistik*, matematiksel yöntemlerle ya da öteki nicel yöntemlerle, özellikle *Yöneylem Araştırması* (Operational Research) yöntemleriyle, sık sık birlikte kullanılır.

*Yöneylem Araştırması*, bir sistemin (istatistik bakışla evrenin) durumunu incelemede, sistemi iyileştirmede, yeni sistemler tasarlama ya da sistem ile ilgili öngörülerde (forecasting) bulunmada, yöneticilerin karar vermelerine yardımcı nicel yöntemlerdir. Yöneylem Araştırmasında yoğunluk karar verme üzerinedir. Seçenek stratejiler oluşturmaya ve bu stratejilerin uygunluğunu test etmeye, performans kriterlerini belirlemeye, sistem davranışını gözlemeye, sistemle ilgili senaryolar hazırlamaya, optimal çözümler elde etmeye yardımcı olur. Her zaman olmasa bile, genelde ileri düzeyde matematik ve istatistik yöntemler gerektirir. Bütçe vb. olanakların, kaynakların dağıtımı; materyallerin seçimi; bekleme hattı, saklama, bilgi erişimi vb. ile ilgili sorunların değerlendirilmesi Yöneylem Araştırması yöntemlerince incelenebilmektedir.

## Çalışmanın Kapsamı ve Kısıtlamaları

Çalışmanın amacı, Kütüphanecilik ve Bilgibilim alanlarında kullanılan istatistik yöntemlerin yoğunluğunu incelemek idi. Wallace(1985), yaptığı çalışmada, istatistik yöntemlerin bu alandaki kullanımını öteki alanlardaki kullanımları ile karşılaştırmıştır. Wallace, yalnızca, ilişki (korelasyon), ki-kare testi, t-testi, varyans çözümlemesi, regresyon çözümlemesi ve konularından oluşan çıkarsamalı istatistik yöntemlerin Kütüphanecilik ve Bilgibilim alanlarındaki kullanım oranının %6, öteki alanlarda ise en az %18 olduğunu belirtmiştir (Wallace, 1985: 406). Kinnucan, Nelson ve Allen(1987) ise bilgilimlerinde kullanılan istatistik yöntemler ile ilgili 'review' çalışmasında daha çok 'Information Science' dergileri üzerinden, iki ya da ikiden çok değişkenli bazı yöntemlerin uygulamaları ile ilgili örnekler vermişlerdir. Kinnucan ve arkadaşları, istatistik yöntemlerin kullanımındaki yaygın bir sorunun, örneğin bir hipotezin testi için uygun bir test bulunabilecek iken bilinen ve basit olan testlerin kullanılması olduğunu; yöntemlerin seçiminde de mantıksal yanılığara düşüldüğünü belirtmişlerdir (Kinnucan, Nelson ve Allen, 1987: 167). Gökkurt(1992: 20), enformetrinin temelinde *Çıkarsamalı İstatistiğin* olduğunu ancak öteki nicel yöntemlerin de bilgiliminde sıkça kullanıldığını vurgulamıştır. Tonta(1998) ise bilimsel araştırmalarda istatistik tekniklerin kullanım ve sunumundaki bazı yanlışlara değinmiştir.

Burada ise, Kütüphanecilik ve Bilgibilim alanlarındaki uygulaması, önceki kesimde verildiği gibi İstatistiğin daha geniş bir alanı üzerinden, ana konulara ve alt konulara göre incelenecektir.

Bu çalışmada istatistik yöntemlerin çeşitliliği ve yoğunluğu üzerinde durulacağından, bir örnekleme yapılmaksızın referans olarak kullanımı yüksek dergiler seçilmiş ve bu dergilerdeki makalelerde kullanılan istatistik yöntemler taranmıştır. Taramada, SSCI(Social Science Citation Index)'de "Information Science and Library Science" kategorisinde bulunan dergiler içinden, 1991-97 arası her dergi için etki faktörlerinin toplamları hesaplanmıştır (ISI-JCR,1992-1998). Çalışmaya toplamları en yüksek 10 dergi alınmak istenmiştir. 1991-97 arası bu kategorideki çoğu derginin etki faktörleri ve toplama göre sıralaması Çizelge 1'de gösterilmiştir. Ancak, bu dergilerden Annual Review of Information Science and Technology (ARIST), yılda bir kez çıktığından ve bir 'review' dergisi olduğundan; Bulletin of the Medical Library Association (B.Med.Lib.Assoc.), yalnızca tıp kütüphanecilerine yönelik özel bir alanda olduğundan; Library and Information Science (Lib. Inf. Sci.) ise bazı makalelerin Japonca olması nedenlerinden çalışmaya alınmamışlardır. Çalışmaya, JASIS (Journal of the American Society for Information Science), J.Doc. (Journal of Documentation), Coll.Res.Lib. (C&R Lib., College and Research Libraries), Libr.Quart (Library Quarterly), Inform. Process. Manag. (IP&M, Information Processing and Management), Libr. Resour. Tech. Ser. (LR&TS, Library Resources and Technical Services), Scientometrics (Sciento.) dergilerinin Hacettepe Üniversitesi Beytepe Kampusu Kütüphanesi, Bilkent Üniversitesi Kütüphanesi ve Ulakbim Cahit Arf Bilgi Merkezi'nde bulunabilen sayıları alınmıştır. Bu dergilere ilişkin incelenen yıllar ve bazı bilgiler Çizelge 2' de verilmiştir. Çalışmanın amacı için, 307 sayı üzerinden 1864 makalenin yeterli olabileceği düşünülmüştür.

Makalelerin taranması sırasında gözlenmiştir ki istatistik yöntemler, genelde, makale adlarında, özetlerde ya da anahtar sözcüklerde geçmemektedir. Bu nedenle, bir makalede istatistik yöntemlerin kullanılıp kullanılmadığı, kullanılmış ise hangi ana ve alt yöntemlerinin kullanıldığı ancak tüm makale satır satır gözden geçirilerek incelenebilmiştir.

Taramada, "editorial", "letter", "notes", "perspective on...", "news" gibi kısa makale türleri göz önüne alınmamıştır.

Ayrıca, amaç, istatistik yöntemlerin yanlış ya da eksik kullanımları olmadığından, tarama sırasında, istatistik yöntemlerin o makale içerisinde doğru olarak kullanılıp kullanılmadığına bakılmamıştır.

İstatistik konuların sınıflandırmasında, "Statistical Abstracts"larda yer alan 'Classification Scheme' kaynak alınmakla birlikte, önceki çalışmalarda kullanılan kategoriler ile Kütüphanecilik ve Bilgibilimi alanlarında İstatistik konularının kullanım sıklıkları da dikkate alınmıştır.

**Çizelge 1. "Information Science & Library Science" Dergilerinin "İmpact Factor" leri**

Sıra Dergi-Adı	IF1991	IF1992	IF1993	IF1994	IF1995	IF1996	IF1997(91-97)
1 ARIST	1,12	1,53	1,00	0,94	1,11	1,53	1,00 8,23
2 JASIS	0,96	1,01	1,07	1,07	1,16	1,23	1,26 7,76
3 J. Doc.	1,07	1,00	1,20	1,03	0,93	1,14	1,25 7,62
4 Coll. Res. Libr.	1,23	1,49	1,16	1,31	0,87	0,19	0,79 7,03
5 Libr. Quart.	0,50	0,77	0,47	0,93	0,97	0,48	1,36 5,49
6 Inform. Proc. Manag.	0,89	0,80	0,62	0,67	0,58	0,62	0,58 4,76
7 B. Med. Libr. Assoc.	0,48	0,48	0,89	0,90	0,67	0,77	0,57 4,76
8 Libr. Ressour. Tech. Ser.	0,81	1,27	1,05	0,36	0,47	0,39	0,40 4,75
9 Libr. Inform. Sci.	***	1,71	0,89	0,53	1,00	0,31	*** 4,44
10 Scientometrics	0,72	0,63	0,52	0,59	0,44	0,58	0,69 4,18
11 J. Acad. Libr.	0,83	0,58	0,66	0,83	0,44	0,18	0,21 3,73
12 Telecommun. Policy.	0,40	0,43	0,49	0,38	0,69	0,43	0,59 3,41
13 Libr. Inform. Sci. Res.	0,35	0,55	0,32	0,47	0,59	0,50	0,59 3,37
14 Online	0,61	0,56	0,56	0,59	0,25	0,42	0,29 3,27
15 Libr. Acquis. Pract. Th.	0,63	0,85	0,73	0,26	0,21	0,34	0,20 3,22
16 Libr. J.	0,52	0,57	0,64	0,35	0,53	0,35	0,25 3,21
17 Int. J. Geogr. Inf. Sci.	***	***	***	0,72	0,72	0,63	1,09 3,16
18 Interlend. Doc. Supply	0,76	0,50	0,47	0,33	0,24	0,42	0,26 2,99
19 Inform. Manage.	***	***	***	0,81	0,83	0,64	0,70 2,98
20 J. Inform. Sci.	0,51	0,49	0,44	0,22	0,47	0,40	0,43 2,97
21 Program-Electron Lib.	0,35	0,41	0,29	0,59	0,48	0,28	0,39 2,78
22 Database	0,43	0,53	0,32	0,56	0,38	0,24	0,26 2,72
23 RQ	0,46	0,48	0,38	0,40	0,30	0,35	0,30 2,65
24 Knowl.Organisation	0,54	0,57	0,38	0,25	0,13	0,18	0,12 2,17
25 Libr. Trends	0,28	0,24	0,28	0,34	0,21	0,40	0,39 2,13
26 Inform. Technol. Libr.	0,40	0,43	0,34	0,26	0,16	0,33	0,19 2,12
27 Gov. Inform. Q.	0,14	0,24	0,53	0,19	0,20	0,27	0,44 2,02
28 On-line CD-ROM Rev.	0,90	0,12	0,10	0,15	0,33	0,22	0,15 1,97
29 Law Libr. J.	0,48	0,21	0,22	0,15	0,30	0,17	0,40 1,92
30 Z. Bibl. Bibl.	0,31	0,33	0,23	0,38	0,21	0,25	0,19 1,90
31 Libri	0,21	0,18	0,31	0,26	0,26	0,33	0,26 1,80
32 Int. J. Inform. Manage.	0,16	0,19	0,16	0,24	0,20	0,38	0,42 1,75
33 Electron. Libr.	0,25	0,27	0,22	0,38	0,17	0,15	0,16 1,59
34 Aslib. Proc.	0,30	0,11	0,13	0,28	0,25	0,10	0,21 1,38
35 Amer. Archivist.	0,40	0,29	0,31	0,08	0,24	0,02	*** 1,33
36 Can. J. Inform. Sci.	0,19	0,07	0,19	0,30	0,13	0,19	0,25 1,31
37 Inform. Outlook	0,10	0,13	0,09	0,33	0,23	0,21	0,21 1,30
38 J. Libr. Inf. Sci.	0,24	0,28	***	0,18	0,28	0,14	0,11 1,22
39 Soc. Sci. Inform.	0,08	0,10	0,13	0,05	0,22	0,27	0,20 1,05
40 P. Asis. Annu. Meet.	0,30	0,10	0,24	0,10	0,22	0,04	0,01 1,01
41 Behav. Soc. Sci. Libr.	0,10	0,12	0,01	0,20	0,31	0,01	0,06 0,81
42 Int. Inf. Libr. Rev.	***	***	0,15	0,04	0,07	0,04	*** 0,31
Toplam	19,0	20,6	18,2	19,0	18,5	16,1	17,2 128,5
Aritmetik Ort.	0,50	0,53	0,47	0,45	0,44	0,38	0,44 3,06
Ortanca	0,44	0,48	0,38	0,36	0,30	0,33	0,30 2,75
Geometrik Ort.	0,41	0,39	0,35	0,35	0,35	0,27	0,32 2,51



Çizelge 2. Taramaya Alınan Dergilere İlişkin Bazı Bilgiler

Dergi	Yıllar	Taranan Dergi Sayısı	Makale Sayısı	Dergi Başına Ortalama Makale Sayısı	Makale Başına Ortalama Referans Sayısı*	Makale Başına Ortalama Yazar Sayısı*	Makale Başına Ortalama Sayfa Sayısı
JASIS	1992-99	78	483	6.2	27	1.7	12
J. Doc.	1990-99	36	159	4.4	24	1.7	18
C&R Lib.	1990-99	56	348	6.2	17	1.5	12
Lib.Quart.	1990-99	36	120	3.3	30	1.4	20
IP&M	1992-99	38	314	8.3	24	1.6	14
LR&TS	1990-99	35	214	6.1	17	1.3	17
Sciento.	1992-99	28	226	8.1	17	1.7	20
<b>Toplam</b>		<b>307</b>	<b>1864</b>	<b>6.1</b>	<b>26</b>	<b>1.6</b>	<b>16</b>

\* Geometrik Ortalama

### Yöntemlerin Dergilere Göre Dağılımları

Tarama sırasında, her bir makale için İstatistik yöntemlerin varlığı ve varsa hangi yöntemler oldukları; ayrıca, eğer ağırlıklı olarak kullanılmış ise matematiksel ve öteki nicel yöntemlerin (özellikle Yöneylem Araştırması) de varlıkları fikir vermesi açısından sayılmıştır. Böylece, çalışmanın ilk aşamasında aşağıdaki kategoriler oluşturulmuştur:

“*İstatistik Yok*”: Taranan makalede betimsel ya da çıkarsamalı anlamda istatistiksel ya da nicel bir yöntem kullanılmamıştır. Makalede, bazı basit ve sık kullanılan oranlar, grafikler, çizelgeler yer alsa bile ya da istatistiksel yöntemler yalnızca kavramsal olarak geçmiş ise o makale bu kategori içinde sayılmıştır.

“*Betimsel*”: Yalnızca *Betimsel İstatistik* yöntemlerin kullanımını gösterir. Ancak buradaki betimsel istatistik terimi, daha önce “*İstatistiksel Süreç*”te verileden daha dar anlamda kullanılmıştır. Bir makalede eğer tek değişkenli düzenleme ve özetleme konuları ile çapraz sıklık çizelgeleri ve grafikler gibi konular uygulanmış ise makale bu kategori içinde sayılmıştır.

“*İleri İstatistik*”: “*İstatistiksel Süreç*”te verilen *Çıkarsamalı İstatistiğin* yanı sıra iki ya da ikiden çok değişkenli istatistiksel yöntemlerin ve olasılık ile ilgili konuların yer aldığı betimsel istatistik yöntemlerini kapsar.

“*Yöneylem*”: Yöneylem Araştırması yöntemlerini kapsar. Eğer bir makalede ağırlıklı olarak bu yöntemler uygulanmış ise makale bu kategoride sayılmış; ileri istatistik yöntemler ağırlıklı olarak kullanılmış ise “*İleri İstatistik*” kategorisinde sayılmıştır.

“*Matematik*”: Matematiksel yöntemler. Eğer bir makalede ağırlıklı olarak matematiksel yöntemler kullanılmış ise makale bu kategoride sayılmıştır.

Kütüphanecilik ve Bilgibilim alanlarında İstatistiğin, özellikle İleri İstatistiğin kullanımını görmek üzere 1864 makalenin genel tarama sonucu Çizelge 3 ve 4’de verilmiştir. Çizelge 3’de, tüm yöntemlerin dergilere göre sayısal, Çizelge 4’de ise yüzdesel dağılışı gösterilmiştir. Çizelgelerde “Sayı” sütunu, ilgili dergiden tarama yapılan dergi sayısını göstermektedir.

**Çizelge 3. Makalelerde Kullanılan Nicel Yöntemlerin Dergilere Göre Sıklık (Frekans) Dağılışı**

Dergi	Sayı	İst.yok	Betimsel	İleri İst.	Yöney	Mat.	Toplam
JASIS	78	125	119	199	11	29	483
J Doc.	36	65	48	34	6	6	159
C&R Lib.	56	127	146	75			348
Lib.Quart	36	65	31	23	1		120
IP&M	38	78	59	126	8	43	314
LR&TS	35	103	72	34	1	4	214
Sciento.	28	16	95	107	1	7	226
<b>Toplam</b>	<b>307</b>	<b>579</b>	<b>570</b>	<b>598</b>	<b>28</b>	<b>89</b>	<b>1864</b>

**Çizelge 4. Makalelerde Kullanılan Nicel Yöntemlerin Dergilere Göre Yüzde Dağılışı**

Dergi	Sayı	İst.yok	Betimsel	İleri İst.	Yöney	Mat.	Toplam
JASIS	78	25.8	24.6	41.2	2.3	6.1	100
J Doc.	36	40.9	30.1	21.4	3.8	3.8	100
C&R Lib.	56	36.5	41.9	21.6			100
Lib.Quart	36	54.2	25.8	19.2	0.8		100
IP&M	38	24.8	18.8	40.2	2.5	13.7	100
LR&TS	35	48.1	33.6	15.9	0.5	1.9	100
Sciento.	28	7.1	42.1	47.4	0.4	3.1	100
<b>Toplam</b>	<b>201</b>	<b>31.1</b>	<b>30.6</b>	<b>32.1</b>	<b>1.5</b>	<b>4.7</b>	<b>100</b>

Çizelge 4'den "İleri İstatistik" yöntemlerinin genel kullanımı %32.1 'dir. Bir başka deyişle makalelerin yaklaşık üçte birinde ileri istatistik yöntemlere gerek duyulmuştur. "Betimsel" den "Matematik" e tüm nicel yöntemlerin kullanıldığı makale oranı ise %68.9'dur; üçte ikiden fazladır.

Dergiler bazında bakıldığında, JASIS, Inform.Process.Manag., ve Scientometrics'de ileri İstatistik yöntemleri kullanan makalelerin oranları %40'ın üzerindedir. Bu dergilerden Scientometrics'in oranı %50'ye yakın, hatta "İstatistik yok" kategorisi dışındaki kategorilerin toplam kullanım oranı %92.9'dur. Kütüphanecilik ve Bilgibilim alanlarında matematiksel ve istatistiksel yöntemlerin uygulanması ile ilgili makaleler yayınlamayı amaçlayan Scientometrics'in amacı da dikkate alındığında, bu dergiye "Econometrica", "Biometrica", "Psychometrica" vb. dergiler gibi, SCI'deki "Statistics" kategorisinde de yer verilmesinin uygun olacağı düşünülebilir.

Matematik yöntemlerin kullanımı, Inform.Process.Manag. ve JASIS'de yoğunlaşırken; JASIS, J.Doc. ve Inform.Process.Manag. dergilerinde Yöneylem Araştırması yöntemlerinin kullanıldığı makale oranları öteki dergilere göre daha yüksektir.

### İleri İstatistik Yöntemlerin Ana Konulara Göre Dağılımları

Yukarıda değinilen ileri istatistik yöntemlerinin, Kütüphanecilik ve Bilgibilim alanlarında kullanımları da göz önüne alınarak, yaklaşık bir sınıflama aşağıda verilmiştir. Çizelge 5'de, bu sınıflamaya ilişkin, makalelerde kullanılan istatistiksel yöntemlerin ana konulara göre sayıları (sıklık dağılışı) ve yüzdeleri sıklık büyüklüğüne göre gösterilmiştir. Sınıflama, uygulanan yöntemlerle ilgili fikir oluşturması açısından önemlidir.

**Çizelge 5. Makalelerde Kullanılan İleri İstatistik Yöntemlerinin Ana Konulara Dağılımı**

Ana Konular	Kullanım Sıklığı	%	Sıra
Olasılık Dağılımları	99	16.6	1
Çok Değişkenli İst. Yönt.	78	13.1	2
Olasılık ve Stokastik Süreçler	73	12.2	3
İlişki ve Birliktelik	71	11.9	4
t-Testleri	63	10.5	5
Regresyon	60	10.0	6
Ki-Kare Testleri	57	9.5	7
Varyans Çöz. ve Deneysel Tas.	39	6.5	8
Kestirim	22	3.7	9
Parametrik Olm. Testler	20	3.3	10
Genel	16	2.7	11
<b>Toplam</b>	<b>598</b>	<b>100.0</b>	

**Çizelge 6. Taranan Makalelerde Kullanılan Ana Konuların Dergilere Göre Yüzde Dağılımı**

Ana Konu	JASIS	Dergiler					
		J. Doc. Lib.	C&R Quart.	Lib.	IP&M	LR&TSSciento.	
Ol. Dağılımları	16.4	16.7	2.7		25.8	5.6	23.4
Çok Değişkenli	14.9	11.1	11.8	27.4	5.0		21.5
Olasılık ve Stok.	13.9	16.7			24.1		9.3
İlişki ve Birlik.	10.0	27.8	13.2	13.6	6.7	19.3	12.1
t-Testi	10.9	5.6	27.6	9.1	5.0	16.7	3.7
Regresyon	8.0		11.8	18.2	5.0	5.6	21.5
Ki-Kare Testleri	7.4	19.3	25.0	9.1	5.8	8.3	3.7
Varyans Çöz.	10.5		3.9	4.5	8.3	5.6	1.9
Kestirim	1.5		1.4		5.1	27.8	1.9
Par.Olm.Testler	5.0			4.5	6.7	2.8	
Genel	1.5	2.8	2.6	13.6	2.5	8.3	1.0
<b>Toplam</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>

Çizelge 6'den görülür ki alışlagelen yöntemler, ilişki ve birliktelik incelemeleri, t-testleri, Ki-Kare testleri tüm dergilerde kullanılmakla birlikte bazı dergilerde daha yoğunlardır. Olasılık dağılımları konusu, özellikle, IP&M, Scientometrics, JASIS ve J. Doc. dergilerinde; Çok Değişkenli İstatistiksel Yöntemler, Lib. Quart., Scientometrics ve JASIS dergilerinde; Olasılık ve Stokastik Süreçler ise IP&M, JASIS ve J. Doc. dergilerinde; Regresyon Çözümlemesi, Scientometrics ve Lib. Quart. dergilerinde sıkça kullanılmıştır.

### İleri İstatistik Yöntemlerin Alt Konu Sıralamaları

Taranan makalelerde rastlanan ileri istatistik yöntemlerine ilişkin tüm tekniklerin ilgili ana konulara dağıtılarak "Alt Konular" adı altında listeleri yapılmaya çalışılmıştır. Her ana konuya ilişkin tüm alt konular makalelerde kullanılma sıklıklarına göre sıralanmışlardır; en sık kullanılanlar, çizelgede, ilk sıralarda yazılmıştır. Bu çizelgelerde geçen teknikler, genelde, doğrudan üzerinde çalışılan ya da uygulama yapılan konuları kapsar. Bir makalede birden çok alt konuya başvurulmuş ise, bunlardan her biri, ilgili ana konu içerisinde yer almıştır.

**Olasılık Dağılımları**

	<b>Sıra</b>
<b>Lotka*</b>	<b>1</b>
Zipf*	2
Bradford*	3
Poisson; Kaydırılmış, Kesilmiş Poisson	4
Genel (Çarpık, Hiperbolik, Çok Tepeli vb.)	5
Leimkuhler, Gen. Leimkuhler*	6
Lognormal	7
Pareto, Gen. Pareto	7
Negative Binom; Kaydırılmış, Kesilmiş N.B.	7
Üstel (Negatif Üstel)	8
Geometrik, Kesilmiş Geometrik	9
Ters-Gaussian Poisson, Gen.Ters-Gaussian P., Kaydırılmış Ters-Gaussian P.	9
Simon, Yule, Schill dağılımları*	10
Hipergeometrik	11
Normal	11
Waring; Gen. Waring, Kaydırılmış Waring*	11
Binom	12
Beta	13
Gamma	14
Öteki dağılımlar (Multinomial, Weibull, Uniform, Bernoulli, Gumbel, Laplace)	

Olasılık dağılımlarına ilişkin alt konularda ilk sırada yer alanlar bibliometrik ve/ya da enformetrik dağılımlardır (\* ile işaretlenenler). Bu dağılımlar, genelde, bir metinde geçen sözcüklerin ortaya çıkış sıklıkları (Zipf); farklı dergiler üzerinden makalelerin saçılımı (Bradford); bilimsel makaleler türünden yazarların verimliliği (Lotka) üzerine verilen yaklaşık dağılımlardır. Leimkuhler, Simon vb. dağılımlar bunların tam olasılık dağılımlarıdır. Bu dağılımlar 1990'larda web sayfaları için atıf sıklıkları üzerinden incelenmiştir (Rousseau,1997).

Bibliometri ve enformetri kavramları, kütüphane modellemesini, bilgi akışındaki nicel beklentileri, bilgi üretim örüntüsündeki (pattern) düzgünlükleri matematiksel ve istatistiksel yöntemler yardımıyla inceler. Burada bibliometri ya da enformetri kavramına, ayrı bir çalışma olması nedeniyle, girilmemiştir.

“Genel dağılımlar”, özellikle, bibliometrik dağılım özelliklerinin tartışıldığı ya da kuramsal bir temele oturtulmaya çalışıldığı genel makaleleri kapsar. Poisson, Pareto, Normal, Lognormal gibi dağılımlar İstatistik'te temel

dağılımları oluştururken; Kaydırılmış Poisson, Kesilmiş Poisson gibi dağılımlar yeniden düzenlenmiş (modifiye edilmiş) dağılımlardır. Bu denli çok çeşitli dağılım olmasının nedeni, çalışılan alanlara bağlı olarak aşırı çarpık dağılımların kuyruklarında görülen farklılıklardır. Örneğin lognormal dağılım, dağılımın genel biçimine iyi bir uyum sağlarken, Pareto ya da Zipf dağılımın uçlarında iyi bir uyum gösterebilir. Varyansın ortalamaya göre çok büyük olduğu durumlarda Poisson ve Gamma dağılımlarının karışımı olan Negatif Binom önerilirken; bu dağılıma seçenek olarak, çok uzun kuyruklu dağılımlarda Ters-Gaussian Poisson dağılımının kullanımı son yıllarda öne çıkmıştır (Sichel, 1992).

<b>Olasılık</b>	<b>Sıra</b>
Olasılık Ölçüleri / Olasılık Hesaplama	1
Entropi	2
Beklenen Değer	3
Koşullu Olasılık	4
Karar Kuramı	5
Bayes Teo.	6
Merkezsiz Limit Yasası	7
Öteki Modeller (Rasch Modeli, Egghe-Duality Modeli, Mansfield Modeli, Shannon-Winer Modeli vb.)	

Olasılık konusu içinde, özellikle olasılıksal arama (searching) çalışmalarını nedeniyle, olasılık hesaplamalarına geniş yer verilmektedir. Böyle bir aramada sorgulamaya uygunluk derecesi atfedilmiş olur.  $i$ . sorgulama verildiğinde  $j$ . belgenin ya da başlığın ortaya çıkma derecesi  $P(j/i) = P(i \cap j) / P(i)$  koşullu olasılığı ile verilebilir ki eşitlik, aynı zamanda, Bayes formülü olarak bilinir. Bireylerin öznel yargıları da  $P(i)$  önsel olasılığı ile hesaplamalara katılabildiğinden, karar verme sürecinde bu olasılık eşitliklerinin sıkça kullanıldığı görülmektedir. Salton ve McGill(1983: 95), bilgi erişim sorununu, üç temel parametrelili bir karar süreci (Karar Kuramı) olarak tanımlamışlardır. Bu parametreler, anlamlılık (relevance) olasılığı, anlamlı kayda erişememe riski ve anlamlı olmayan kayda erişme riskidir.

Öte yandan Shannon'un bilgi kuramına göre bir sözcüğün ortaya çıkış olasılığı arttıkça içerdiği bilgi azalır. Buradan,  $p_k$  örneğin bir simgenin ortaya çıkış olasılığı iken, *entropi* adı verilen bir belgedeki ortalama bilgi miktarı,  $-\sum_k p_k \log(p_k)$  ile tanımlanır. Benzer olarak, bir  $i$ . belgedeki  $k$ . terimin önemi ya da  $i$ . ve  $j$ . belgelerdeki  $k$ . özellikler arasındaki bilgi farkı vb. olasılıklar ya da bu olasılıklara dayalı ağırlıklar hesaplanabilir. Daha ileri düzeyde, ör-

neğin Rasch modeli gibi n. dergiye i. yılda atıf yapılma olasılığı gibi olasılıkların hesaplanması için, modeller de geliştirilmiştir (Lasee,1992; Alvarez and Pulgarin,1996).

<b>Bibliometrik Süreçler</b>	<b>Sıra</b>
Markov Zinciri	1
Waring Süreci	2
Markov Poisson Modeli	3
Poisson Süreci	4
Gamma-Poisson Süreci	5
Homojen Olm.Doğum Süreci	6

Bilginin elde edilmesi, saklanması, iletilmesi vb. çalışmaların, zamanla ilişkisi olması sonucu stokastik süreçler konularının, zaman zaman "Bibliometrik Süreçler" adı altında sıkça yer aldığı görülmektedir. Bu süreçler, bilginin zaman üzerinden sergilenen değişimini modellerken; Simon'un bibliometrik dağılımlar üzerindeki çalışmasıyla ağırlık kazanmıştır (Simon, 1955).

<b>İlişki ve Birliktelik Katsayıları</b>	<b>Sıra</b>
Pearson İlişki Kats.ve Testleri	1
Spearman rho İlişki Kats.	2
Cramer V Kats.	3
Kısmi İlişki Kats.	4
Kendall Tau Kats.	5
Phi-Kats.	6
Goodman-Kruskal Gamma Kats.	7
Kappa Kats.	8
Öteki konular (Nokta iki-dizili, Dört düzeyli, Konkordans Kats., Ardışık İlişki, Çapraz İlişki, Geometrik Yak. vb.)	

İlişki ve birliktelik katsayıları (correlation , association), değişkenler arasındaki ilişkinin uzunluğunu verirler. Nicel değişkenler için Pearson ilişki katsayısı, sıralanmış veriler için Pho katsayısı kullanılır. 2x2 çapraz sıklık çizelgeleri için Phi katsayısı; r satır s sütunlu çizelgeler için Cramer-V katsayısı vb. kullanılabilir. Değişkenlerden bazıları sabit tutulduğunda iki değişken arasındaki ilişki, kısmi ilişki katsayısı olarak adlandırılır. Kappa katsayısı, örneğin iki kodlama sistemi arasındaki, uyumun derecesinin bir ölçüsü olarak verilir. Öte yandan Pearson ilişki katsayısı, geometrik olarak iki vektör arasındaki açının kosinüsü olarak da, özellikle indekstemede, sıkça kullanılmıştır.

<b>Çok Değişkenli İstatistiksel Yöntemler</b>	<b>Sıra</b>
Kümeleme Çöz. (Cluster Analysis)	1
Etken Çöz. (Factor Analysis)	2
Çok Boyutlu Ölçekleme (Multidimensional Scaling-MDS)	3
Uygunluk Çöz. (Correspondence Analysis)	4
Çok Değişkenli Varyans Çöz.(Mutivar.An.of Variance-MANOVA)	5
Ayrırcı Çöz. (Discriminant Analysis)	6
Kanonik Kor. (Canonical Correlation)	7
Temel Bileşenler Çöz. (Principle Component Analysis)	7
Çokdeğ.Tek.Ölç.Varyans Çöz. (Multivar.Repeated Meas.Anova)	8

Bu yöntemler, çok sayıda gözlem üzerinden ikiden çok değişken arasındaki ilişkileri ya da bağımlılık yapısını incelemek amacıyla kullanılır. En sık kullanılan ilk dört yöntem, boyut küçültme teknikleridir;<sup>3</sup> değişkenler arasındaki karmaşık ilişkilerin basitleştirilmesine yöneliktir(Tatlıdil,1996: 138). *Kümeleme*, belgelerin, terimlerin ya da yazarların sınıflandırılmasına istatistiksel bir yaklaşımdır. Amaç, değişkenler yani özellikler göz önüne alınarak, gözlemleri benzerliklerine göre gruplandırmak; böylece, araştırmacıya uygun özet bilgi sağlamaktır. *Etken Çözümlemesi*, kümelemenin yaptığı değişkenler bazında yapar. Örneğin her bir yazarın her bir dergiden aldıkları atıf sayıları incelenirken, gözlemler yazarlardan, değişkenler dergilerden oluşsun; eğer yazarlar sınıflandırılacaksa kümeleme çözümlemesi, eğer dergiler gruplandırılacaksa etken çözümlemesi kullanılmış olacaktır. Çok boyutlu ölçekleme ise kümeleme yöntemine benzer; gözlemler arasındaki uzaklık, yakınlık ya da benzerlik öğeleri kullanılarak, veri kümesinin bir “resmi” ya da bir “haritası” çıkartılmaya çalışılır. Aynı amaçla, eğer değişkenlerde nitel özellikler varsa (kategorik veriler), MDS yerine uygunluk çözümlemesi uygulanır. Öteki yöntemler çok sayıda bağımsız değişkenle, özellikle birden çok bağımlı değişken arasındaki bağımlılık yapısının incelenmesi ile ilgilidir.

<b>Regresyon Çözümlemesi</b>	<b>Sıra</b>
Basit Doğrusal Reg.	1
Çoklu Doğrusal Reg.	2
Dönüşümler	3
Logistik Reg.	4
Değişken Seçimi	5
Doğrusal Olm. Regresyon (Gompertz Modeli, Büyüme Eğrileri,vb.)	6
Çokterimli (Polynomial) Reg.	7
Üstel Reg.	8
Değişen Varyanslılık Sorunu	9



Ters Regresyon (Calibration)	10
Bayeşçi Reg.	10
Poisson Reg.	10
Genelleştirilmiş EKK	11
Negative Binom Reg.	11
Sağlam Reg. Kestirimi	11

Regresyon (Bağlanım) çözümlemesi, bağımsız değişkenler ile tek bağımlı değişken arasındaki bağıntının çözümlemesidir. Regresyon çözümlemesinde kullanılan tekniklerin çeşitliliği dikkat çekicidir. Basit ve çoklu doğrusal regresyon, bağımlı değişkenin kitap yaşı ya da uzunluğu gibi nicel ve sürekli olmasını gerektirir; gözlemlerin bağımsızlığı gibi varsayımları vardır. Genelde, bağımlı değişkenin bilinmeyen değerlerinin ya da ileride gözlenmesi beklenen değerlerin önceden kestirimi amaçlanır. Eğer ilişkiler doğrusal değilse, varsayımlar sağlanmıyorsa, veri türleri farklı ise çözümleme için bir yol, veriler üzerinde logaritmik, karekök gibi dönüşümler yapmaktır. Eğer bağımlı değişken iki düzeyli nitel verilerden oluşmuş ise logistik regresyon, gösterdiği dağılım biçimlerine göre Üstel, Poisson, Negative Binom vb. regresyon türleri kullanılabilir. Kütüphanelerin büyüme modellerini görmek için doğrusal olmayan regresyon, çözümlemede önsel olasılıklar ya da öznel bilgiler kullanılmak isteniyorsa Bayeşçi regresyon uygulanabilir.

Kütüphanede günlük, aylık vb. dolaşım (ödünç vermeler) gibi, İnternet'deki bilgilerin günlük, haftalık değişimleri gibi zamana göre elde edilecek verilerin incelenmesinde *Zaman Dizileri* çözümlemesi kullanılabilir. Zaman dizilerinde, bir zaman döneminde olan şeyler bir sonraki zaman dönemini de etkiler. İleriki zamanlara öngörülerde (forecasting) bulunabilmek için verilerin zaman dizileri teknikleri ile ele alınması gerekir. Zaman dizileri çözümlemesi, zamana bağlı veri yapılarından ötürü, çoğu kez, *Stokastik Süreçler* konusu ile beraber incelenir.

Varyans Çözümlemesi ve Deneysel Tasarım	Sıra
Anova (bir yönlü)	1
Deneysel Tasarım (2 <sup>k</sup> Etkensel, Split Plot, Mixed, vb.)	2
Çoklu Karşılaştırmalar (Scheffe testi, Tukey testi, vb.)	3
Anova (iki yönlü)	4
Tekrarlı Ölçümlü Anova	5
Varyans Homojenliği (Levene, Cochran C testleri)	6
Anova (üç yönlü)	7
Anova(dört yönlü)	8
Kovaryans Çözümlemesi	8

Özellikle, İnternetin yaygınlaşması ile birlikte genişleyen uygulama alanları bazı etkenlerin araştırmacı denetiminde *Deneyisel Tasarım* yoluyla elde edilecek veriler üzerinden incelenmesini gerekli kılmaktadır. Bağımlı değişken sürekli iken, kategorik bağımsız değişken düzeyleri arasındaki farklılıklar, etkileşimler varyans çözümlemesi (Analysis of Variance-Anova) ile incelenebilir. Bu çalışmalarda, denetlenemeyen değişkenlerin etkilerini azaltmak için, rasgeleleştirme esasına dayalı deneysel tasarım yöntemleri tercih edilir. Böylece daha az gözlemle daha düşük bir standart hata elde edilebilir. Değişken yani etken sayısına göre bir yönlü, iki, üç, vb. yönlü çözümler yapılabilir. Varyansların benzer olduğu durumlarda hangi grupların birbirinden farklı olduğu çoklu karşılaştırmalar ile test edilebilir (Hicks, 1994). Aynı birim üzerinden bağımlı gözlemler elde edilmiş ise, bağımlılık etkisini azaltmak ya da zaman etkisini ortadan kaldırmak üzere *tekrarlı ölçümlü varyans çözümlemesi* kullanılabilir.

### Student t-Testleri

### Sıra

t-testi (ortalama)	1
t-testi (ortalamalar arası fark)	2
t-testi (yüzde)	3
z-testi	4
t-testi (yüzdeler arası fark)	5
t-testi (eşler arası)	6

Student t-testleri, varyans çözümlemesinin özel durumlarıdır. Eğer yalnızca iki düzeyli tek bir değişken varsa, bu düzeylerin ya da grupların ortalamaları ya da oranları arasındaki farklılıkların testi, t-testleri ile; yeterince büyük örneklem için z-testi ile yapılır.

### Ki-Kare Testleri

### Sıra

Ki-Kare (Bağımsızlık testi)	1
Ki-Kare (Uyum iyiliği testleri)	2
Log-Doğrusal Model	3
Fisher Exact testi	4

Ki-Kare testleri, uygulamada, özellikle anket değerlendirmede çok sık rastlanan bu testler, iki değişkenli çapraz sıklık ilişkisiyi test etmede uygundur. Örneklemden elde edilen bir sıklık dağılımının bilinen kuramsal bir dağılıma uygunluğu da ki-kare yöntemi ile test edilir. Öte yandan, çok boyutlu çizelgelerde yani değişken sayısının ikiden çok olduğu bu tür çizelgelerde, birlikteliklerin incelenmesi için, ki-kare yerine Log-doğrusal modellerin kullanılması bu alanda da giderek yaygınlaşmaktadır.

<b>İstatistiksel Kestirim</b>	<b>Sıra</b>
Güven Aralığı (yüzde)	1
Bayesçi Çıkarsama	2
Parametre Kest. (genel)	3
EÇO Kestirimi	4
Güven Aralığı (ortalama)	4
Olabilirlik Fonk.	4
Bootstrap Kest.	5
Sağlamlılık	5
Yanılıgı Kareler Ort. (MSE)	5
Güven Aralığı (yüzdeler arası)	6
Öteki (Bonferroni Güven Ar., Çekirdek (Kernel) Kest., Konvülüsyon Yaklaşımı, Önkestirim (Prediction), Yanlılık)	

İstatistikte ortalama, oran gibi her örnekleme değeri bir nokta kestirimi olarak adlandırılır. Bu kestirimler, En Çok Olabilirlik (EÇO), En Küçük Kareler (EKK), Bayesçi Yöntemler gibi yollardan elde edilirler. Örneğin Lotka parametrelerinin kestiriminde EKK yöntemi kullanılır. Dağılımların bilinip bilinmemesine göre, verilerin özelliklerine göre kestirim yöntemleri değişecektir. Uygulamada bu yöntemlerin adları geçmese de bilinen eşitlikler yardımıyla bu değerler elde edilmişlerdir. Bu nedenle doğrudan bu konular üzerinde yapılmış çalışmalar, çok azdır. Bununla beraber, belli bir yanılma payı ya da riski altında nokta kestirimi üzerinden evrendeğerlerin hangi sınırlar arasında olabileceği de güven aralıkları yoluyla incelenebilir. Varsayımların sağlanmaması, aykırıdeğerlerin varlığı vb. durumlarda değişik kavramlar ortaya çıkar.

<b>Parametrik Olmayan Testler</b>	<b>Sıra</b>
Kolmogorov-Smirnov testi (Uyum İyiliği)	1
Wilcoxon İşaretli-Sıra (Signed-rank) testi	2
Mann-Whitney U testi	3
Wilcoxon Eşli İşaretli-Sıra testi (Matched-Paired Signed-Ranks)	4
G-Kare (Uyum İyiliği) testi	5
Kruskal-Wallis testi	5
İşaret testi (Sign test)	5
Cochran Q testi	6

Öteki (Cramer-von Misses; Binom, Friedman, Lehmann, Poisson testleri)

*Parametrik Olmayan Testler*, varyans çözümlemesi, t-testleri gibi parametrik testler adı verilen testler de varsayımların sağlanmadığı durumlara karşılık gelen testlerdir. Özellikle uyum iyiliğine karşılık gelen testler sınıflama gerektirmediğinden, çoğu kez, ki-kareye tercih edilirler. Tek örneklem testlerinde ortalama yerine, genelde, ortanca kullanılır. Uygulama, evrenin simetrik olup olmadığına göre değişir. İki örneklem arasındaki farklılığın örnekleme rasgeleliğinden gelip gelmediğinin testi, İşaret testleri, U-testi gibi testlerle yapılırken; örneğin Kruskal-Wallis testi ile k sayıda örneklem arasındaki farklılık ya da bağımsızlık incelenir. Örneklemelerin yayılma ya da değişkenlik yönünden farklılıkları da Q testi gibi testlerle araştırılır.

### **Genel Yöntemler**

Uygulama (Ekonomiden: Gini İndeksi, Lorenz Eğrisi, Cost-Benefit vb.)

Benzetim (Simulation)

Örnekleme (Tabakalı, Sistematik Örk., Nitelik (Attribute) Örk., Fussler Tekniği)

Uzaysal (Spatial) Veri Çöz.

Kalite Denetimi (Quality Control)

Meta Çözümlemesi

Öte yandan, genel yöntemlere ilişkin ekonomi alanından alınarak, özellikle yoğunlaşma ve eşitsizlik kavramları için Gini katsayısının uygulaması bu alanda oldukça yaygındır. Lorenz eğrisi nüfusa karşı gelir dağılımını ifade eder. Gini katsayısı ya da İndex'i, Lorenz eğrisi ile köşegen arasındaki alanın toplam alana oranı olarak ifade edilir ve eşitsizliğin derecesini gösterir (Rao, 1988: 242). Kütüphanecilik ve Bilgi Bilimleri alanında, özellikle, atıf, yazar, makale vb. yoğunluk derecelerini görmek amacıyla ele alınmaktadır.

Benzetim yöntemleri, İstatistik'de, bilgisayarda denetlenebilir bir örnekleme türüdür. Değişik parametre kestirimlerinin, vb. sonuçların, evren ya da uzun dönem bilgileri önceden verilerek, belli dağılımlara dayalı rasgele sayılardan elde edilen yapay veriler üzerinden karşılaştırılmalarını sağlar. Benzetim, bir sistemin örneğin bir kütüphane sisteminin sorunlarının incelenmesi, iyileştirilmesi, sistem ile ilgili çözüm stratejilerinin karşılaştırılabilmesi gibi durumlarda, genelde Yöneylem Araştırması yöntemleri içerisinde de yer alan ve sıkça kullanılan bir yöntemdir. Örneklem üzerinden tüm incelemelerde örnekleme yöntemleri kullanılmış olsa da doğrudan bu örnekleme yöntemlerinin incelendiği çalışma sayısı oldukça azdır. Ancak örneğin kütüphanelerde yapılan örneklemede ortaya çıkabilecek yanlılığı ortadan kaldırmak için Fuzzler tekniği gibi özel yöntemlerin kullanıldığı çalışmalar vardır (Nickerson, 1992).

Öte yandan, jeolojik ve biyolojik verilerde olduğu gibi tek bir değişkenin kendi içindeki ilişkilerin incelendiği *Uzaysal Çözümleme*; İstatistik ve Örneklem yöntemlerinin sıklıkla uygulandığı, örneğin, verimliliğin ya da bilginin kalitesinin ölçülmek istendiği *Kalite Denetimi*, doğrudan olmasa bile öteki konular ile birlikte incelenen konulardır. Farklı yer ve zamanlarda ancak aynı amaç ve konularla ilgili yapılmış çalışmaların nasıl birleştirilebileceği, farklı çalışmaların değişkenleri arasındaki ilişkilerin nasıl elde edilebileceği vb. sorunların çözümü için verilen *Meta Çözümlemesi* yöntemi ise disiplinler arası bir inceleme olarak bilinir.

### **Yöneylem Araştırması**

Genetik Algoritmaları

Karar Destek Sistemleri (Decision Support Systems)

Uzman Sistemler (Expert Systems)

Eniyileme (Optimizasyon)

Tam Sayılı Programlama

Ağ (Network) Çözümlemesi

Dinamik Programlama

Doğrusal Programlama

Matematiksel Programlama

Burada esas amaç olmamakla birlikte, Kütüphanecilik ve Bilgibilimleri alanında kullanılan öteki nicel yöntemlere de kısaca değinelim. En sık geçen yöntemlerden *Genetik Algoritmalar*, doğadaki biyolojik evrimin işleyiş biçiminin bilgisayar yoluyla taklidi olan bir *Stokastik Optimizasyon* yöntemidir. *Karar Destek Sistemleri*, yönetici yargısının ön planda olduğu, örneğin bir kütüphanenin gelecekteki etkinliği ile ilgili senaryolar kurma, senaryoları geliştirme, karşılaştırma, karar verme çalışmalarını kapsar. *Uzman Sistemler* konuları da genelde, karar destek sistemleriyle birlikte ele alınan konulardır. Bunlar, zaman zaman İstatistik yöntemlerle birlikte ya da bağımsız olarak kullanılırlar.

### **Matematik Yöntemler**

Boole Cebri

Bulanık (Fuzzy) Kümeler

Vektör Uzayları

Fractal Geometri

Topoloji

Mandelbrot Parametreleri

Matrisler Kuramı

Ölçme Kuramı

## SVD Ayrışımı

Diferansiyel Denklemler vb.

Ayrıca bazı matematiksel yöntemler de daha önce verildiği gibi JASIS, IP&M, Scientometrics ya da J. Doc. dergilerinde, istatistik yöntemlerle, özellikle olasılık kavramlarıyla birlikte ya da bağımsız olarak, giderek artan sıklıkta uygulanmaktadır. Örneğin, belge erişiminde, Boolean Modeli, Olasılıksal Model, Vektör Uzayı Modelleri, Bulanık Kümeler Modeli ve TIRS (topolojik bilgi erişim) Modeli gibi kavramlar kullanılır (Everett and Cater,1992). Boolean Modelleri, Boole cebri yani “0-1” ya da “var-yok” mantığına göre işlerken; bulanık küme kavramında birimlerin belli bir derece ile küme içinde olup olmadıkları mantığı geçerlidir. Belli bir desenin sonsuz kez tekrarlayarak ortaya çıkardığı şekilleri konu alan Fractal Geometri ve öteki matematiksel kavramlar, ileri düzeyde matematik bilgisi gerektirdiğinden bu alan içindeki kullanımlarının azlığı doğaldır.

## Sonuç

1. 1991-97 arası etki faktörü en yüksek olan 7 dergideki çalışmaların yaklaşık %70'i nicel yöntemleri kullanmaktadır.
2. Çalışmaların %32'si ileri istatistik yöntemlerini kullanmaktadır. Bazı dergilerde ise %40'ın üstünde kullanım vardır. Scientometrics, aynı zamanda bir istatistik dergisi gibi işlev görmektedir.
3. Öyle görülüyor ki *Dağılımlar Kuramı, Stokastik Süreçler ve Çok Değişkenli İstatistik Yöntemler* ileriki zamanlarda Kütüphanecilik ve Bilgibilim alanlarında daha sık kullanılır duruma gelecektir.
4. Geleneksel Ki-Kare testleri, t-testleri ise giderek, özellikle Bilgibilimlerinde, kullanımı azalan yöntemler olmaktadır. Ki-kare yerine üç ya da daha çok boyutlu çizelgelerin incelenmesi ve log-doğrusal modellerin kullanımı giderek yaygınlaşabilecektir.
5. Bu çalışmalara bakarak Kütüphanecilik ve Bilgibilim alanlarında, özellikle hangi konuların temel olarak İstatistik derslerinde verilebileceği belirlenebilir. İstatistik dersinde ağırlık uygulamalı konularda olduğundan ve lisans düzeyinde temel bir eğitim verilmeye çalışıldığından ileri konular işlenmemektedir. Bununla birlikte, Temel İstatistik derslerinde okutulan, “*Verilerin Düzenlenmesi, Özetlenmesi, Oranlar, İki Değişkenli Veriler, İlişki Kavramı ve Katsayıları, Basit Olasılık Hesaplamaları ve Dağılımları, Örneklem, İstatistiksel Kestirim ve Güven Aralıkları, Bazı Hipotez Testleri, İlişki ve Bağlanım (Regresyon) Kavramı, Bibliometrik İncelemeler*” konuları yanında,

## İleri Olasılık Kavramları ve Modelleri

Temel Olasılık Dağılımları ve Bibliometrik Dağılımlar

Temel Stokastik Süreçler

Aralık ya da Oran Ölçekli Olmayan Veriler için İlişki Katsayıları

Sıkça Kullanılan Çok Değişkenli Yöntemler

Çoklu ve Doğrusal Olmayan Bazı Regresyon Kavramları

Bazı Deneysel Tasarım Kavramları

Bazı Log-Doğrusal Model Yöntemleri

konularının da bu derslerde ya da daha ileri düzey derslerde en azından kavramsal olarak verilmesinin yararlı olacağı söylenebilir.

**KAYNAKÇA**

- Alvarez, P. and A. Pulgarin. (1996) "The rasch model. measuring information from keywords: The diabetes fiel", JASIS, 47 (6): 468-476.
- Blalock, H.M. (1979) *Social statistics*, 2<sup>nd</sup> Ed., New York: McGraw-Hill.
- Bookstein, A. (1988) "Applications of the bibliometric distributions", *informetrics* 87/88, 5-13, L.Egghe and R. Rousseau (Editors), Amsterdam: Elsevier Science Publishers.
- Everett, D.M. and S.C. Cater. (1992) "Topology of document retrieval systems", JASIS, 17 (2): 658-673.
- Gökkurt, Özlem. (1992) "Enformetrinin istatistiksel temeli: Vardamlı (inferential) istatistik teknikler", *Türk Kütüphaneciliği*, 6 (1): 18-21.
- Hicks C. R. (1994) *Deney düzenlemede istatistiksel yöntemler*, Çev.: Z. Muluk, Ö. Toktamış, S. Kurt, E. Karaağaoğlu, İzmir: Ege Üniversitesi.
- Feeney, M. And M. Grieves. (1994) *The Value and impact of information*, Bowker-Sour Ltd.
- Hafner A.W. (1998) *Descriptive statistical techniques for librarians*, 2<sup>nd</sup> Ed., Chicago: American Library Association.
- Houser, L.J. and G.J. Lazorick. (1978) "Introducing a significant statistics component into a library science resarch methods course", *J.of Education for Librarianship*, 18:182-88
- Institute for Scientific Information (1992,....,1998) 1991,....,1997 Social Science Citation Index Journal Citation Reports: A Bibliometric analysis of social science journals in the ISI
- Database, Eugene Gardfield (Editor), Philadelphia.
- Kinnucan, M.T., M.J. Nelson, B.L. Allen. (1987) "Statistical methods in information science research", *ARIST*, 22, 147-178.

- Lasee, R.M. (1992) "A Gray code based ordering for documents on shelves classification for browsing and retrieval", *JASIS*, 43 (4): 312-322.
- Nickerson M. (1992) "pH: Only a pierce of the preservation puzzle: A Comparison of the preservation studies at Brigham Young, Yale and Syracuse Universities", *Libr.Res.Tech.Ser.*, 36 (1): 105-112.
- Pao, M. Lee (1989) *Concepts of information retrieval*, Englewood: Libraries Unlimited, Inc.
- Pritchard, A. (1969) "Statistical bibliography or bibliometrics", *Journal of Documentation*, 25 (4): 348-349.
- Rao, I.K.Ravichandra. (1988) "Probability distributions and inequality measures for analysis of circulation data", *Informetrics* 87/88, 231-247, L.Egghe and R. Rousseau(Editors), Amsterdam: Elsevier Science Publishers.
- Rousseau R. (1997) "Sitations: An Exploratory study", *Cybermetrics*, 1 (1): 1-8
- Salton, G. and M.J. McGill. (1983) *Introduction to modern information retrieval*, New York: McGraw-Hill.
- Sichel, H.S. (1992) "Note on a strongly unimodal bibliometric size frequency distribution", *JASIS*, 43 (4): 299-303.
- Simon, H.A. (1955) "On a class of skew distribution functions", *Biometrika*, 42: 425-440.
- Tatlıdil, Hüseyin. (1996) *Uygulamalı çok değişkenli istatistiksel analiz*, Ankara: Akademi Matbaası.
- Tonta, Yaşar. (1999) "Bilimsel araştırmalarda istatistik tekniklerin kullanımı ve bulguların sunumu üzerine", *Türk Kütüphaneciliği*, 13 (2): 112-124.
- Wallace D.P. (1985) "The Use of statistical methods in library and information science", *JASIS*, 36 (6): 402-410.
- Wyllys, R.E. (1978) "Teaching descriptive and inferential statistics in library schools", *Journal of Education for Librarianship*, 19: 3-20.