

T.C.
PAMUKKALE ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

**KONTENJANS TABLOLARINDA LOG-LİNEER
ANALİZ İLE HOMOJENİTE ANALİZİNİN
KARŞILAŞTIRILMASI**

Beyza AKDAĞ

BIYOİSTATİSTİK ANABİLİM DALI
YÜKSEKLİSANS TEZİ

88980

DANIŞMAN
Yrd. Doç. Dr. Mevlüt TÜRE

Şubat-1998
DENİZLİ
T.C. YÜKSEKÖĞRETİM KURULU
DOKÜMANTASYON MERKEZİ

Pamukkale Üniversitesi Sağlık Bilimleri Enstitüsü
Biyoistatistik Yüksek Lisans Programı
çerçevesinde yürütülmüş olan bu çalışma, aşağıdaki jüri tarafından
Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 20/03/1998

JÜRİ BAŞKANI
Prof. Dr. Kazım ÖZDAMAR

ÜYE
Doç. Dr. Mehmet BOSTANCI

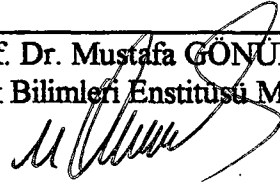
ÜYE
Doç. Dr. İzzettin HATİP

ÜYE
Doç. Dr. Diler ASLAN

ÜYE
Yrd. Doç. Dr. Mevlüt TÜRE

Sağlık Bilimleri Enstitüsü Yönetim Kurulu'nun ...14..04..1998..... gün ve
.....98/02-1..... sayılı kararıyla onaylanmıştır.

Prof. Dr. Mustafa GÖNÜLLÜ
Sağlık Bilimleri Enstitüsü Müdürü



İÇİNDEKİLER

	<u>Sayfa</u>
İÇİNDEKİLER	iii
TABLolar DİZİNİ	v
ŞEKİLLER DİZİNİ	vi
SİMGELER VE KISALTMALAR DİZİNİ	vii
GİRİŞ	9
GENEL BİLGİLER	12
1. İKİ VE ÇOK YÖNLÜ KONTENJANS TABLOLARI	12
2. LOG-LİNEER ANALİZ	12
2.1. Genel LLA	14
2.2. Lojit LLA	14
2.3. Aşamalı LLA	14
3. İKİ YÖNLÜ TABLOLARDA LOG-LİNEER ANALİZ	14
4. ÇOK YÖNLÜ TABLOLARDA LOG-LİNEER ANALİZ	16
5. UYUM ANALİZİ	17
5.1. Uzaklık Ölçüsü	18
5.2. Ana Eksen ve Koordinatların Hesaplanması	20
5.3. \mathcal{R}^p 'deki Analiz, Faktörlerin Hesaplanması	21
5.4. \mathcal{R}^n 'deki Analiz ile İlişkiler	22
5.5. En İyi Eş Zamanlı İzdüşümün Bulunması	24
5.6. Açıklanma Derecesi ve Kare Korelasyonların Hesaplanması	25
5.7. S Matrisinin Simetrikleştirilmesi	31
6. HOMOJENİTE ANALİZİ	33
6.1. Z ile bağlantılı Burt tablosu	34
6.2. İki Değişkenli Uyum	35
6.3. Birinci Eşitlik (1 ve 2 arasındaki)	35
6.4. İkinci Eşitlik (1 ve 3 arasındaki)	37

	<u>Sayfa</u>
6.5. İki'den Daha Fazla Değişken İçin Genelleme	38
6.6. Çoklu Analizlerin Özellikleri	41
6.7. ϕ_q Bileşenlerinin Merkezinin Belirlenmesi	41
6.8. Bir Kategori ve Bir Değişkene Uygun Varyansın Oranı	42
6.9. \mathcal{R}^n 'deki p Kategorinin Konfigürasyonunun Boyutlandırılması	43
6.10. En İyi Eşzamanlı Gösterim	43
7. GRAFİKSEL GÖSTERİM	44
GEREÇ VE YÖNTEM	45
1. VERİLER	45
2. ANALİZ YÖNTEMİ	45
BULGULAR	46
1. GERİYE DOĞRU AŞAMALI LOG-LİNEER ANALİZİN SONUÇLARI	46
2. EN UYGUN LOG-LİNEER MODELİN ETKİLEŞİMLERİNİN HOMOJENİTE ANALİZİ	50
2.1. ADY*EAA*COO Etkileşimi İçin HA	50
2.2. BDY*EAA*COO Etkileşimi İçin HA	51
2.3. ADY*BDY*BED Etkileşimi İçin HA	52
2.4. AED*EAA Etkileşimi İçin HA	53
2.5. AED*BED Etkileşimi İçin HA	54
2.6. ADY*AED Etkileşimi İçin HA	55
TARTIŞMA VE SONUÇ	57
ÖZET	60
SUMMARY	61
KAYNAKLAR DİZİNİ	62
TEŞEKKÜR	65
ÖZGEÇMİŞ	66

TABLolar DİZİNİ

Tablo-1: 2*2 biçimindeki kontenjans tablosunun genel yapısı.	14
Tablo-2: 2*2 biçimindeki kontenjans tablosunun logaritmik değerleri.	15
Tablo-3: k yönlü etkilerin LR Ki-kare istatistik sonuçları.	46
Tablo-4: En uygun model bulunmadan önce ilgili etkileşimlere ait serbestlik dereceleri, kısmi ki-kare değerleri ve önemlilik düzeyleri.	46
Tablo-5: En uygun modelin etkileşimlerinin LR ki-kare sonuçları.	48
Tablo-6: En uygun modelin, en uygun model bulunmadan önceki kısmi ilişkilerinin (etkileşimler) standart hata, z ve p değerleri.	49
Tablo-7: ADY*EAA*COO etkileşimine ilişkin çoklu çapraz tablo.	50
Tablo-8: BDY*EAA*COO etkileşimine ilişkin çoklu çapraz tablo.	51
Tablo-9: ADY*BDY*BED etkileşimine ilişkin çoklu çapraz tablo.	52
Tablo-10: AED*EAA etkileşimine ilişkin çapraz tablo.	53
Tablo-11: AED*BED etkileşimine ilişkin çapraz tablo.	54
Tablo-12: ADY*AED etkileşimine ilişkin çapraz tablo.	55

ŞEKİLLER DİZİNİ

Şekil-1: j'inci eleman ile α ekseninin kare korelasyonu.	27
Şekil-2: ADY*EAA*COO etkileşimi için HA.	50
Şekil-3: BDY*EAA*COO etkileşimi için HA.	52
Şekil-4: ADY*BDY*BED etkileşimi için HA.	53
Şekil-5: AED*EAA etkileşimi için HA.	54
Şekil-6: AED*BED etkileşimi için HA.	55
Şekil-7: ADY*AED etkileşimi için HA.	56



SİMGELER VE KISALTMALAR DİZİNİ

μ	Genel Ortalama
X	Kategorik değişken
Y	Kategorik değişken
X_1	X değişkeninin birinci alt kategorisi
X_2	X değişkeninin ikinci alt kategorisi
Y_1	Y değişkeninin birinci alt kategorisi
Y_2	Y değişkeninin ikinci alt kategorisi
\bar{X}_i	Satırların gözlenmiş frekanslarının ortalamaları
\bar{Y}_j	Sütunların gözlenmiş frekanslarının ortalamaları
f_{ij}	Kontenjans tablosundaki gözelerin frekansı
n_{ij}	Göze frekanslarının logaritmik değerleri
α_x	X'in ana etkisi
α_y	Y'nin ana etkisi
α_{xy}	X ve Y'nin etkileşimi
K	Kontenjans tablosu matrisinin elemanları
F	Oran matrisi
D_p	Satır marjinal olasılıkları diagonal matrisi
D_n	Sütun marjinal olasılıkları diagonal matrisi
\mathcal{R}^p	p boyutlu öklid uzayı
\mathcal{R}^n	n boyutlu öklid uzayı
X	Veri matrisi
M	Simetrik uzaklık matrisi
N	Diagonal matris
u	Birim vektör
\hat{v}	u eksenini üzerindeki n noktanın görüntüsü
λ	Özdeğer
ϕ	Faktör
ψ	λ özdeğerli bir faktör
p	Cevap kategorilerinin toplam sayısı
Q	Değişken sayısı
q	Herhangi bir değişken
p_q	Herhangi bir değişkenin aldığı değer
Z	n satırlı, p sütunlu matris
R	Z matrisinin özeti
B	Z ile ilişkili Burt'ün kontenjans tablosu
D	Diagonal matris
I	Birim matris
γ_q	Linear alt uzayı
HA	Homojenite Analizi

SD Serbestlik Derecesi
S Simetrik olmayan herhangi bir matris
A:(ADY) Annenin Doğum Yeri
B:(AED) Annenin Eğitim Durumu
C:(EAA) Eşler Arasında Akraba Evliliği
D:(BDY) Babanın Doğum Yeri
E:(COO) Çocukta Özur Olup Olmadığı
F:(BED) Babanın Eğitim Durumu



GİRİŞ

Günümüzde özellikle tıp ve biyoloji alanlarında birimlerin çok sayıda değişkenlerine ilişkin veriler elde edilir. Bu verilerin analizinde çok değişkenli analizlerin uygulanması objektif kararların alınmasını sağlar. Son yıllarda bilgisayar teknolojisinin hızla gelişmesi ve buna paralel olarak istatistik paket programlarının kullanımının oldukça yaygın ve kolay hale gelmesi nedeniyle çok değişkenli istatistiksel analiz yöntemlerinden yararlanmak oldukça yaygın hale gelmiştir (1,2).

Çok değişkenli yöntemlerden yaygın olarak kullanılanları; Çok Değişkenli Varyans Analizi, (Multivariate Analysis of Variance, MANOVA), Ayırma Analizi (Discriminant Analysis, DA), Faktör Analizi (Factor Analysis, FA), Temel Bileşenler Analizi (Principal Component Analysis, PCA), Çok Boyutlu Ölçekleme Analizi (Multidimensional Scaling, MDS), Kümeleme Analizi (Cluster Analysis, CLA), Log-lineer Analiz (Log-linear Analysis, LLA), Kanonik Korelasyon Analizi (Canonical Correlation Analysis, CCA), Uyum Analizi (Correspondence Analysis, CA), Homojenite Analizi (Homogeneity Analysis, HA) olarak sayılabilir (3,4).

Araştırma sonucu elde edilen verilerin sayım ya da ölçüm yoluyla elde edilmesine göre kullanılan istatistiksel metotlar değişiklik göstermektedir. Eğer veriler ölçüm yoluyla elde edilmiş ise bu verilerin analizinde parametrik, sayım yoluyla elde edilmiş ise parametrik olmayan istatistiksel yöntemler kullanılır. Sayım yoluyla elde edilen verileri analiz etmek için genellikle kontenjans tabloları analizi kullanılır. Kontenjans tablolarında söz konusu olan değişkenler arasındaki bağımsızlığın araştırılmasında son yıllara kadar ki-kare bağımsızlık analizi kullanılmıştır. Ki-kare bağımsızlık ve türdeşlik analizlerinde sadece söz konusu ilgili değişkenler arasında bağımsızlık olup olmadığına karar verilir. Bağımlılık olması durumunda ise satır ve sütun sayısının çok olduğu durumlarda bağımlılığın ölçüsü hakkında açıklayıcı bilgiler elde etmek ve elde edilen sonuçları yorumlamak oldukça zordur. Bu nedenle, çok değişkenli kontenjans tablolarının analizinde, satır ve sütun etkileşimlerinin belirlenerek değerinin hesaplanmasında ve yorumların yapılmasında ki-kare analizleri yetersiz kalmaktadır (5,6,7,8).

Çok değişkenli analiz yöntemlerinden LLA, HA, lineer olmayan temel bileşenler analizi, lineer olmayan kanonik korelasyon analizi ve çok boyutlu ölçekleme analizi ki-

kare analizinin kullanıldığı her türlü tablonun yanında, özellikle değişken sayısının ikiden fazla olduğu çok değişkenli kontenjans tablolarında değişkenler arasındaki karmaşık yapıda olan ilişkilerin daha kolay anlaşılabilmesine imkan veren analiz yöntemleridir (9,10,11,12).

LLA yöntemi, kontenjans tablolarında değişkenler arasındaki birlikteliğin ve etkileşimin yapısal özelliklerini modeller ile ortaya koyan bir yöntemdir. Uyum analizi ve onun genişletilmiş şekli olan HA, veri matrislerinin satır ve sütun değişkenlerinin ve alt kategorilerinin birbirleri ile bağımlılığını grafiksel gösterimle ortaya koyan yöntemlerdir. Uyum analizi iki değişken söz konusu olduğu durumlarda, HA ise ikiden daha fazla değişken olduğu durumlarda kullanılırlar.

Yapılan çalışmalarda CA ve HA diğer çok değişkenli istatistiksel yöntemler ile karşılaştırılmıştır;

Goodman (1986), CA modellerinin nasıl genel uygunluk analiz yaklaşım yönteminin genişletilmiş hali olarak sunulabileceğini ve log-bilineer modellerin nasıl genel log-lineer model yaklaşımının genişletilmiş hali olacağını açıklamıştır (13).

de Leeuw (1988), iki yönlü tablolarda bağımsız olan modellerden hata terimlerinin analizinde faydalı olabilen CA tekniğini sunmuştur. Uyum analizinin özellikle satır ve sütun sayılarının çok fazla olduğu büyük tablolar için uygun bir analiz yöntemi olduğunu belirtmiştir (14).

van der Heijden ve Worsley (1988) ile van der Heijden ve ark. (1989), CA ile LLA yöntemini karşılaştırmışlar ve bu iki yöntemin birbirinin tamamlayıcısı olduğunu savunmuşlardır (15).

van der Heijden ve ark. (1989) CA ile LLA yöntemini karşılaştırmışlar ve bunun sonucu olarak CA ve onun genişletilmiş şekli olan HA'nın log-lineer modellerde artıkların analizi için kullanılabilirliğini belirtmişlerdir (16).

Tenenhaus ve Young (1985), yapmış oldukları bir çalışmada kategorik çok değişkenli verileri ölçmek için çeşitli metotları tartışmışlar, bunların arasında HA'nın çoğu kez önerilen bir teknik olduğunu ve kategorik değişkenli değişken matrisiyle bir nesnenin analizi için uygun bir yöntem olabileceğini belirtmişlerdir (17).

van der Burg ve ark. (1988), yapmış oldukları bir arařtırmada HA'nın deęiřkenlerin sayısının homojenlięini maksimize etmede kullanılan bir metod olduęundan ve deęiřkenlerin her birinin kategorilerinin transformasyonlarını ya da ölçümlerini belirleyen bir yöntem olduęundan bahsetmişlerdir (18).

Bu çalışmada; HA ve LLA yönteminin birbirinin tamamlayıcısı olup olmadıklarını belirlemek amacıyla; 1996 yılında Denizli ili merkezinde yaşayan 1000 aile üzerinde yapılan bir arařtırmada, akraba evlilięi yapmaya ve çocuęun özürlü olmasına neden olan deęiřkenlerin belirlenmesi için “annenin eęitim durumu”, “annenin doğum yeri”, “babanın eęitim durumu”, “babanın doğum yeri”, “eřler arasında akraba evlilięi” ve “çocuklarında özür olup-olmadığı” deęiřkenlerine geriye doğru aşamalı LLA yöntemi ile HA uygulanmıştır.

GENEL BİLGİLER

1. İKİ VE ÇOK YÖNLÜ KONTENJANS TABLOLARI

Değişkenler, ele alınan ölçü yöntemlerine göre ya da çeşitli ölçütlere göre iki ya da daha fazla gruba ayrılarak sınıflandırılabilirler. Özellikle niteliksel değişkenlerin aynı anda olanaklı özellikler taşıdığı birden fazla grupta yer almasının sözkonusu olduğu koşullarda, sayısal olarak değeri saptanan değişkenleri, iki ya da daha çok ölçüte göre sınıflandırarak R*C tabloları biçiminde ele almak gerekir. Bu tür tablolara "kontenjans tablosu", "çapraz tablo" ya da "R*C tablosu" gibi isimler verilmektedir (8,19,20,21,22,23).

Örnekleme yer alan birimin 2 ya da daha fazla değişken üzerinde aynı anda sınıflandırılmasıyla oluşturulan tabloya çok değişkenli çapraz tablo adı verilir. İstatistik birimlerinin aynı anda gözlenen iki ya da daha çok değişken arasında bir ilişkinin olup olmadığını ortaya koymak için çapraz tablolardan yararlanır. Değişkenler sınıflayıcı, sıralayıcı, aralıklı ve oransal ölçümlere göre incelenmektedirler. Araştırma kapsamına giren birimler sınıflayıcı ve sıralayıcı ya da gruplanarak kategorik hale dönüştürülen aralıklı, oransal ölçekle ölçülmüş değişkenlere göre tablolaştırılmaktadır. Tablolarda yer alan sıklıklar belli özelliklerini gösteren birimlerin sayılarını ifade etmek üzere, her bir değişkenin tablonun bir boyutu ile ifade edildiği çok boyutlu çapraz tablo yapısı ile verilmektedir. Çapraz tablonun iki boyutlu olması durumunda satır ve sütun değişkenleri arasındaki homojenliği test etmek için Pearson'un χ^2 istatistiği ve G^2 olabilirlik-oran istatistiği (Likelihood-Ratio, LR) kullanılmaktadır. Birçok durumda, bir sonuca etki eden faktör sayısı ikiden fazladır. Bu nedenle yaygın olarak karşılaşılan çapraz tablolar çok boyutludur (9,10,21,23,24).

Kontenjans tablolarıyla özetlenen kategorik yapıya sahip değişkenler arasındaki ilişki ve etkileşimler, log-lineer modellerin ilgili terimleri yolu ile ifade edilirler. Bu modeller yardımıyla söz konusu olan etkilerin ve etkileşimlerin anlamlı olup olmadığını belirten ön savların sınanmasına imkan tanınmış olur (25).

2. LOG-LİNEER ANALİZ

LLA, R*C biçiminde düzenlenmiş çok yönlü (iki ya da daha fazla) çapraz tablolarda kategorik değişkenler arasındaki birlikteliğin ve etkileşimin yapısal özelliklerini modeller ile ortaya koyan bir yöntemdir (26,27).

LLA, deęişkenler arasındaki iliřkileri olasılık daęılımları yoluyla karakterize eder. Deęişkenler arasında var olan iliřkilerin en uygun matematiksel fonksiyonla temsil edilmesinin arařtırılması özellikle iliřki belirleyici arařtırmalarda çok önemlidir. Kontenjans tablolarıyla özetlenen kategorik deęişkenler arasındaki iliřki ve etkileşimler log-lineer modellerin terimleriyle ifade edilebilirler (23,27,28,29).

Log-lineer modeller, çok yönlü çapraz tablolarda deęişkenler arasındaki karmařık iliřkileri ortaya çıkarmak için faydalıdır. Log-lineer modeller çoklu regresyon modelleri ile benzerdir çünkü bu yöntemde de deęişkenler faktör ve baęımlı deęişken olarak ayrılır, ve deęişkenler arasındaki iliřkiler modeller yoluyla ortaya çıkarılır (4,27,30,31).

İki ya da daha fazla kategorik deęişkenle bir frekans tablosu biçimlendirildięi zaman, deęişkenler arası iliřkiler LLA yardımıyla modellenabilir. LLA modeli beklenen hücre frekansının doğal logaritmasını gösterir (birçok yönde alınmış varyans analizine benzerdir. Çünkü bu analiz yönteminde de deęişkenler faktör ve baęımlı deęişken olarak alınır ve deęişkenler arasındaki etkileşimler ortaya çıkarılır) ve bu durumda model, etkileşimler ve ana etkilerin bir lineer kombinasyonudur (11,27,32,33).

Kategorik verilerin analizinde söz konusu olan kriterlerden birisi de olasılık orandır. Olasılık Oran (Odds Ratio: OR) deęişkenler arasındaki etkileşimi ifade eder. Etkenin varlığında bir olayın ortaya çıkma olasılıęının etken söz konusu olmadığında ortaya çıkma olasılıęına göre kaç kat daha fazla olduęunu gösteren bir istatistiktir (23).

LLA'da çözümlenmeler yapılırken verilerin durumuna göre üç temel çözümlenme yönteminden yararlanılır.

Bu yöntemler;

- Genel LLA
- Lojit LLA
- Ařamalı LLA

olarak adlandırılır.

2.1. Genel LLA

Bu yöntemde R*C biçiminde tablolaştırılmış X ve Y kategorik değişkenleri birer faktör olarak alınır. Gözelerdeki frekanslar ise bağımlı değişken olarak alınır. Tablodaki göze frekansları üzerinde X ve Y'nin ana etkileri ve X ve Y'nin etkileşimi değerlendirilir (23).

2.2. Lojit LLA

Kategorik değişkenlerden birisi bağımlı değişken olarak, geride kalan değişkenler ise bağımsız değişken olarak alınır. Bu modelin uygulanabileceği veri yapılarında bağımlı değişken mutlaka kategorik olmak zorundadır. Bağımsız değişken ise kategorik ya da sürekli olabilir (23).

2.3. Aşamalı LLA

İki kategorik değişken arasındaki ilişki test edileceği zaman, yapılan işlem kontenjans tablosu oluşturmak ve uygun ki-kare istatistiğini hesaplamaktır. Fakat ikiden daha fazla kategorik değişkene sahip olduğunda, bir yaklaşım olarak iki yönlü tabloların serilerini oluşturmak ve her biri için ilgili ki-kare istatistiklerini hesaplamak olacaktır. Bunun yanı sıra bu strateji, değişkenler arasındaki etkileşimler nedeniyle oldukça aksar. Ayrıca yorum da artan değişken sayısına bağlı olarak karmaşıklığa dönüşür. İdeal çözüm, aynı anda değişkenlerin grupları arasındaki etkileşimleri içeren bütün değişkenler arasındaki ilişkileri test etmektir. Bu da aşamalı LLA yöntemi ile yapılabilir (4,23,27).

3. İKİ YÖNLÜ TABLOLARDAN LOG-LİNEER ANALİZ

Kontenjans tablolarının en basit biçimi 2*2 biçimindeki tablolardır. Bu tabloların genel yapısı Tablo-1'de verilmiştir.

Tablo-1: 2*2 biçimindeki kontenjans tablosunun genel yapısı.

X	Y		Toplam
	Y1	Y2	
X1	f_{11}	f_{12}	R1
X2	f_{21}	f_{22}	R2
Toplam	C1	C2	N

Tablo-2: 2*2 biçimindeki kontenjans tablosunun logaritmik değerleri.

X	Y		Toplam
	Y1	Y2	
X1	n_{11}	n_{12}	\bar{X}_1
X2	n_{21}	n_{22}	\bar{X}_2
Toplam	\bar{Y}_1	\bar{Y}_2	μ

$$n_{ij} = \log(f_{ij})$$

olup f_{ij} , kontenjans tablosundaki gözelerin frekansını gösterir. R*C tablolarında satır ve sütunların gözlenmiş frekanslarının ortalamaları (\bar{X}_i ve \bar{Y}_j) ve genel ortalama (μ) sırasıyla;

$$\bar{X}_1 = (n_{11} + n_{12}) / 2 \quad \bar{X}_2 = (n_{21} + n_{22}) / 2 \quad (1)$$

$$\bar{Y}_1 = (n_{11} + n_{21}) / 2 \quad \bar{Y}_2 = (n_{12} + n_{22}) / 2 \quad (2)$$

$$\mu = (n_{11} + n_{12} + n_{21} + n_{22}) / 4 \quad (3)$$

şeklinde hesaplanır.

Eğer X ile Y arasında bağımlılık söz konusu ise bu durumda etkileşim de göz ardı edilemeyeceği için model;

$$\ln(f_{ij}) = \mu + \alpha_x + \alpha_y + \alpha_{xy} \quad (4)$$

şeklinde olur. Bu modele doymuş model adı verilir. X ve Y bağımsız ise bu durumda etkileşim terimi söz konusu olmayacağından model;

$$\ln(f_{ij}) = \mu + \alpha_x + \alpha_y \quad (5)$$

olur. Bu modellerde α_x X'in ana etkisini, α_y Y'nin ana etkisini ve α_{xy} ise X ve Y'nin etkileşimini gösterir ve

$$\alpha_x = \bar{X}_1 - \mu \quad \text{ya da} \quad \alpha_x = \bar{X}_2 - \mu \quad (6)$$

$$\alpha_y = \bar{Y}_1 - \mu \quad \text{ya da} \quad \alpha_y = \bar{Y}_2 - \mu \quad (7)$$

$$\alpha_{xy} = \ln(f_{ij}) + \alpha_x + \alpha_y - \mu \quad (8)$$

olarak hesaplanır.

İki değişken arasındaki birlikteliğin önemliliği normal yaklaşım ile;

$$z = \alpha_{xy} / \sqrt{\text{Var}(\alpha_{xy})} \quad (9)$$

olarak hesaplanır. Burada standart sapma aşağıdaki gibi hesaplanır:

$$\sqrt{\text{Var}(\alpha_{xy})} = \sqrt{1/16[(1/f_{11}) + (1/f_{12}) + (1/f_{21}) + (1/f_{22})]} \quad (10)$$

4. ÇOK YÖNLÜ TABLOLARDA LOG-LİNEER ANALİZ

Çok yönlü R*C tabloları, iki yönlü tabloların genişletilmiş şekli olup doymuş model,

$$\ln(f_{ijk...m}) = \mu + \alpha_X + \alpha_Y + \dots + \alpha_{XY} + \alpha_{YZ} + \dots + \alpha_{XYZ} + \alpha_{XYV} + \dots + \alpha_{XYZ...W} \quad (11)$$

biçiminde ifade edilir.

Kontenjans tablolarında;

$$n_{ijk...m} = \ln(f_{ijk...m}) \quad (12)$$

olup $f_{ijk...m}$ kontenjans tablosundaki gözelerin frekansını gösterir. X, Y, Z, ..., W bağımsız ise bu durumda etkileşim terimi söz konusu olmayacağı için model;

$$\ln(f_{ijk...m}) = \mu + \alpha_X + \alpha_Y + \alpha_Z + \dots + \alpha_W \quad (13)$$

olur.

Üç yönlü R*C*D tablosunda satır ve sütunların gözlenmiş frekanslarının ortalamaları (\bar{X}_i , \bar{Y}_j ve \bar{Z}_k) ve genel ortalama (μ) sırasıyla;

$$\begin{aligned} \bar{X}_i &= \sum_{j=1}^C n_{ijk} / C \\ \bar{Y}_j &= \sum_{i=1}^R n_{ijk} / R \\ \bar{Z}_k &= \sum_{k=1}^D n_{ijk} / D \\ \mu &= \sum_{i=1}^R \sum_{j=1}^C \sum_{k=1}^D n_{ijk} / (R \times C \times D) \end{aligned} \quad (14)$$

şeklinde hesaplanır. Burada, X, Y ve Z satır ve sütunları göstermektedir. Modellerdeki ana etkiler ve etkileşim terimleri,

$$\alpha_x = \bar{X}_x - \mu \quad (15)$$

$$\alpha_{xy} = \ln(f_{ij}) + \alpha_x + \alpha_y - \mu$$

biçiminde, iki değişken arasındaki birlikteliğin önemliliği ise normal yaklaşım ile;

$$z = \alpha_{xy} / \sqrt{\text{Var}(\alpha_{xy})} \quad (16)$$

şeklinde hesaplanır (4,23).

5. UYUM ANALİZİ

Uyum Analizi, çok değişkenli bir analiz yöntemi olup iki veri seti (bir kontenjans tablosunun satır ve sütunları) arasındaki bağımlılığı çok boyutlu inceleyen ve grafiksel olarak gösterimini veren bir yöntemdir. Uyum analizi, kontenjans tablosu şeklindeki verileri noktalar ile resmedilmiş grafiksel gösterime çevirir (8,14,16,32,34,35,36).

Pearson ki-kare analizi, satır ve sütun sayısı az olan kontenjans tablolarının analizinde pratik bir yöntemdir. Bu yöntemde, satır ve sütunları birbirleri ile karşılaştırmak için satır ve sütun yüzdelerine bakılmaktadır. Satır ve sütun yüzdeleri arttıkça satır ve sütunların birbirleri ile tek tek karşılaştırılmaları güçleşmekte hatta bazen imkansızlaşmaktadır. Satır ve sütunların birbirlerinden bağımsız olup olmadıklarını ve hangi satır ya da satırlar ile hangi sütun ya da sütunların türdeşlik gösterdiğini ve bu türdeşliğin ne düzeyde olduğunu belirlemede CA kullanılmaktadır (8,16,32,35).

Uyum Analizinde kullanılan ve temel olan matris, kontenjans tablosu matrisinin (**K**) elemanlarının toplam birim sayısına oranlanarak bulunan **F** matrisidir. Bu matris bir oran matrisidir ve aşağıdaki gibi gösterilir (8,32,37,38,39,40,41).

$$\mathbf{K} = \begin{bmatrix} k_{11} & k_{12} & \dots & k_{1p} \\ k_{21} & k_{22} & \dots & k_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ k_{n1} & k_{n2} & \dots & k_{np} \end{bmatrix}_{n \times p} \quad \mathbf{F} = \begin{bmatrix} f_{11} & f_{12} & \dots & f_{1p} \\ f_{21} & f_{22} & \dots & f_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ f_{n1} & f_{n2} & \dots & f_{np} \end{bmatrix}_{n \times p} \quad (17)$$

Bu matrisin elemanları;

$$f_{ij}=k_{ij}/k ; i \in p, j \in n \quad (k = \sum_{i,j} k_{ij}) \quad (18)$$

olup satır ve sütun marjinal olasılıkları;

$$k_{i.} = \sum_{j=1}^n k_{ij} \quad f_{i.}=k_{i.}/k ; i \in p \quad (19)$$

$$k_{.j} = \sum_{i=1}^p k_{ij} \quad f_{.j}=k_{.j}/k ; i \in n \quad (20)$$

şeklinde yazılır. Bu olasılıkların sıfırdan büyük olduğu varsayılır ($f_{i.}>0, f_{.j}>0$). Satır ve sütun marjinal olasılıkları diagonal matris (**D**) şeklinde aşağıdaki gibi ifade edilir (8,16,32,35,37,38,39).

$$\mathbf{D}_p = \begin{bmatrix} f_{1.} & 0 & \dots & 0 \\ 0 & f_{2.} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & f_{p.} \end{bmatrix}_{p \times p} \quad \mathbf{D}_n = \begin{bmatrix} f_{.1} & 0 & \dots & 0 \\ 0 & f_{.2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & f_{.n} \end{bmatrix}_{n \times n} \quad (21)$$

\mathcal{R}^p 'nin (\mathcal{R}^p : p boyutlu öklid uzayı) i. vektörünün j. bileşeni (satır yüzdesi) ve \mathcal{R}^n 'in (\mathcal{R}^n : n boyutlu öklid uzayı) j. vektörünün i. bileşeni (sütun yüzdesi) aşağıdaki gibi yazılabilir.

$$\left\{ \frac{k_{ij}}{k_{i.}} \right\}_{j=1,2,\dots,p} \quad \left\{ \frac{k_{ij}}{k_{.j}} \right\}_{i=1,2,\dots,n} \quad (22)$$

5.1. Uzaklık Ölçüsü

Uyum Analizinde yalnızca uzaklık türü ölçü söz konusudur. Satır noktaları ve/veya sütun noktaları arasındaki uzaklıklar ki-kare uzaklığı olarak isimlendirilirler (8,14,16,32,34,35,38,39,40,41,42,43,44).

i ve i' gibi iki satır arasındaki uzaklık:

$$d^2(i,i') = \sum_{j=1}^n \frac{1}{f_{.j}} \left(\frac{f_{ij}}{f_{i.}} - \frac{f_{i'j}}{f_{i'.}} \right)^2 \quad (23)$$

olup, benzer olarak, j ve j' gibi iki sütun arasındaki uzaklık:

$$d^2(j, j') = \sum_{i=1}^n \frac{1}{f_i} \left(\frac{f_{ij}}{f_j} - \frac{f_{ij'}}{f_{j'}} \right)^2 \quad (24)$$

şeklinde yazılabilir. Bu şekilde tanımlanan uzaklık ki-kare uzaklığı olarak isimlendirilir.

Ki-kare uzaklığının seçilmesinin nedeni şu şekilde açıklanır;

1. Eğer özdeş satır ya da sütun profillerine sahip olan iki sütun toplanırsa, bu durumda sütunlar arasındaki uzaklık değişmez.

2. Eğer aynı dağılımlı profillere sahip iki satır toplanırsa o zaman satırlar arasındaki uzaklık değişmez.

Eğer basit profiller ile sütunlar toplanırsa o zaman satırların temsil edilmesi çok önemli bir değişikliğe uğramaz. Bir i_0 satırını i_1 ve i_2 satırlarının toplamı şeklinde ifade edersek, o zaman f_{i_0} relatif populasyon frekansı olur ve,

$$f_{i_0} = f_{i_1} + f_{i_2} \quad (25)$$

şeklinde hesaplanır.

j ve j' sütunları arasındaki $d^2(j, j')$ uzaklığı yalnız iki terim içerir. T_1 ve T_2 olarak isimlendirilen bu terimler i_1 ve i_2 kullanılarak aşağıdaki gibi hesaplanırlar.

$$T_1 + T_2 = \frac{1}{f_{i_1}} \left(\frac{f_{i_1j}}{f_j} - \frac{f_{i_1j'}}{f_{j'}} \right)^2 + \frac{1}{f_{i_2}} \left(\frac{f_{i_2j}}{f_j} - \frac{f_{i_2j'}}{f_{j'}} \right)^2 \quad (26)$$

Bu toplam T_0 ile gösterilirse:

$$T_0 = \frac{1}{f_{i_0}} \left(\frac{f_{i_0j}}{f_j} - \frac{f_{i_0j'}}{f_{j'}} \right)^2 \quad (27)$$

olur ve

$$T_0 = T_1 + T_2 = f_{i_0} \left(\frac{f_{i_0j}}{f_{i_0} \cdot f_j} - \frac{f_{i_0j'}}{f_{i_0} \cdot f_{j'}} \right)^2 \quad (28)$$

biçiminde yazılabilir.

5.2. Ana Eksen ve Koordinatların Hesaplanması

X (n,p)'lik bir veri matrisi, M \mathcal{R}^p 'de pozitif tanımlı simetrik uzaklık matrisi ve N 'de (n,n)'lik diagonal matris olsun. N diagonal matrisinin diagonal elemanları da n noktanın ağırlığıdır (38,40).

\mathcal{R}^p 'nin birim vektörü u şu özelliğe sahiptir:

$$u'Mu = 1 \quad (29)$$

u eksenini üzerindeki n noktanın görüntüsü aşağıdaki gibidir.

$$\hat{v} = XMu \quad (30)$$

Görüntülerin (projeksiyonlar) ağırlıklı kareler toplamı :

$$\hat{v}'N\hat{v} \quad (31)$$

olup,

$$u'MX'NXMu$$

olur. Bu kantite;

$$u'Mu = 1 \quad (32)$$

ile maksimize olur. u vektörü ana eksen olarak isimlendirilir ve

$$X'NXMu = \lambda u \quad (33)$$

eşitliği sağlanır.

$$\phi = Mu \quad (34)$$

ile tanımlı olan ve u eksenini üzerindeki projeksiyonlar (izdüşümler) bazen “faktör” olarak isimlendirilirler. Bu da aşağıdaki eşitliği sağlar:

$$MX'NX\phi = \lambda\phi \quad (35)$$

Uzaklık için 1'in bir normuna sahip faktörler M^{-1} 'dir. Ve aşağıdaki gibi ifade edilir;

$$\phi'M^{-1}\phi = u'MM^{-1}Mu = u'Mu = 1 \quad (36)$$

5.3. \mathfrak{R}^p 'deki Analiz, Faktörlerin Hesaplanması

Uzaydaki n nokta, $\mathbf{D}_n^{-1}\mathbf{F}$ 'in n satırıdır. \mathbf{u} , \mathfrak{R}^p 'deki uzaklık için birim vektör olsun.

Böylece

$$\mathbf{u}'\mathbf{D}_p^{-1}\mathbf{u} = 1 \quad (37)$$

ve \mathbf{u} ekseninde n projeksiyonun vektörü

$$\mathbf{D}_n^{-1}\mathbf{F}\mathbf{D}_p^{-1}\mathbf{u} = \hat{\mathbf{v}} \quad (38)$$

olur.

Maksimize edilen kantite, ağırlıklı kareler toplamıdır ve

$$\hat{\mathbf{v}}'\mathbf{D}_n\hat{\mathbf{v}} \quad (39)$$

ya da

$$\mathbf{u}'\mathbf{D}_p^{-1}\mathbf{F}'\mathbf{D}_n^{-1}\mathbf{F}\mathbf{D}_p^{-1}\mathbf{u} \quad (40)$$

olarak yazılır ve

$$\mathbf{u}'\mathbf{D}_p^{-1}\mathbf{u} = 1 \quad (41)$$

olarak bulunur. \mathbf{u} aşağıdaki matrisin özvektörüdür;

$$\mathbf{S} = \mathbf{F}'\mathbf{D}_n^{-1}\mathbf{F}\mathbf{D}_p^{-1} \quad (42)$$

ve en büyük λ özdeğerlidir:

$$\mathbf{S}\mathbf{u} = \lambda\mathbf{u} \quad (43)$$

\mathbf{S} 'nin genel terimleri olan s_{ij} 'ler şu şekilde yazılır;

$$s_{ij} = \sum_{i=1}^n \frac{f_{ij}f_{ij}}{f_i f_j} \quad (44)$$

\mathbf{u} vektörü birinci ana eksendir.

$$\boldsymbol{\varphi} = \mathbf{D}_p^{-1}\mathbf{u} \quad (45)$$

vektörü birinci faktör olarak isimlendirilir.

ϕ faktörü

$$\mathbf{A} = \mathbf{D}_p^{-1} \mathbf{F}' \mathbf{D}_n^{-1} \mathbf{F} \quad (46)$$

matrisinin bir özvektörüdür.

\mathbf{u} ana eksenini üzerindeki n noktanın izdüşümleri aşağıdaki denklemin bileşenleridir;

$$\mathbf{D}_n^{-1} \mathbf{F} \mathbf{D}_p^{-1} \mathbf{u} = \mathbf{D}_n^{-1} \mathbf{F} \phi \quad (47)$$

Çok genel olarak, eğer \mathbf{u}_α , λ_α öz değerine sahip \mathbf{S} matrisinin özvektörü ise (\mathbf{u}_α α 'nci temel eksendir) o zaman $\phi_\alpha = \mathbf{D}_p^{-1} \mathbf{u}_\alpha$, α 'nci faktör olmaktadır ve \mathbf{u}_α eksenini üzerindeki n noktanın projeksiyonu $\mathbf{D}_n^{-1} \mathbf{F} \phi_\alpha$ 'nın bileşenleridir (38,40).

5.4. \mathcal{R}^n 'deki Analiz ile İlişkiler

Aşağıdaki matris \mathcal{R}^n 'de maksimize edilirse;

$$\mathbf{v}' \mathbf{D}_n^{-1} \mathbf{F} \mathbf{D}_p^{-1} \mathbf{F}' \mathbf{D}_n^{-1} \mathbf{v} \quad (\mathbf{v}' \mathbf{D}_n^{-1} \mathbf{v} = 1) \quad (48)$$

\mathbf{v} ana eksenini büyük özdeğeri $\mathbf{F} \mathbf{D}_p^{-1} \mathbf{F}' \mathbf{D}_n^{-1}$ matrisinin özvektörüdür.

(43) nolu eşitlikten hareketle;

$$\mathbf{F}' \mathbf{D}_n^{-1} \mathbf{F} \mathbf{D}_p^{-1} \mathbf{u} = \lambda \mathbf{u} \quad (49)$$

olur ve her iki taraf da $\mathbf{F} \mathbf{D}_p^{-1}$ ile çarpılırsa;

$$\mathbf{F} \mathbf{D}_p^{-1} \mathbf{F}' \mathbf{D}_n^{-1} (\mathbf{F} \mathbf{D}_p^{-1} \mathbf{u}) = \lambda (\mathbf{F} \mathbf{D}_p^{-1} \mathbf{u}) \quad (50)$$

elde edilir. Buradan \mathbf{v} vektörünün $\mathbf{F} \mathbf{D}_p^{-1} \mathbf{u}$ ile orantılı olduğu görülür. \mathbf{D}_n^{-1} nedeniyle $\mathbf{F} \mathbf{D}_p^{-1} \mathbf{u}$ 'nun normu λ 'ya eşittir ve böylece;

$$\mathbf{v}' \mathbf{D}_n^{-1} \mathbf{v} = 1 \quad (51)$$

olup aşağıdaki eşitlikler yazılabilir;

$$\mathbf{v} = \frac{1}{\sqrt{\lambda}} \mathbf{F} \mathbf{D}_p^{-1} \mathbf{u} \quad \mathbf{u} = \frac{1}{\sqrt{\lambda}} \mathbf{F}' \mathbf{D}_n^{-1} \mathbf{v} \quad (52)$$

v eksenini üzerindeki p deęişken-noktaların koordinatları aőaęıdaki denklemin bileőenleridir;

$$\mathbf{D}_p^{-1} \mathbf{F}' \mathbf{D}_n^{-1} \mathbf{v} = \mathbf{D}_p^{-1} \mathbf{F}' \boldsymbol{\psi} \quad (53)$$

burada ($\boldsymbol{\psi} = \mathbf{D}_n^{-1} \mathbf{v}$) olmaktadır.

ψ , λ özdeęerli bir faktör olarak isimlendirilir. Bu temel v eksenini üzerindeki izdüőüm-iőlemcisidir. Buna göre aőaęıdaki formüller yazılabilir;

$$\boldsymbol{\psi} = \mathbf{D}_n^{-1} \mathbf{v} = \frac{1}{\sqrt{\lambda}} \mathbf{D}_n^{-1} \mathbf{F} \mathbf{D}_p^{-1} \mathbf{u} = \frac{1}{\sqrt{\lambda}} \mathbf{D}_n^{-1} \mathbf{F} \boldsymbol{\phi} \quad (54)$$

$$\boldsymbol{\phi} = \mathbf{D}_p^{-1} \mathbf{u} = \frac{1}{\sqrt{\lambda}} \mathbf{D}_p^{-1} \mathbf{F}' \mathbf{D}_n^{-1} \mathbf{v} = \frac{1}{\sqrt{\lambda}} \mathbf{D}_p^{-1} \mathbf{F}' \boldsymbol{\psi} \quad (55)$$

Bu iki iliőki seti; uzaydaki bir ana eksen üzerindeki noktaların koordinatlarının aynı özdeęere sahip dięer uzayın faktörünün bileőenlerine oranını gösterir.

Bu sonuçlar bir λ_α öz deęerine sahip bir birim için genellenebilir: ϕ_α , ψ_α ve λ_α için iliőkiler aynı őekildedir.

Böylece faktörler arasındaki őu iliőkiler bulunur:

$$\psi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} \mathbf{D}_n^{-1} \mathbf{F} \phi_\alpha \quad (56)$$

$$\phi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} \mathbf{D}_p^{-1} \mathbf{F}' \psi_\alpha \quad (57)$$

α eksenini üzerindeki koordinatların vektörleri $\hat{\phi}_\alpha$ ve $\hat{\psi}_\alpha$ olarak isimlendirilir;

$$\hat{\phi}_\alpha = \phi_\alpha \sqrt{\lambda_\alpha} \quad (58)$$

$$\hat{\psi}_\alpha = \psi_\alpha \sqrt{\lambda_\alpha} \quad (59)$$

(56) ve (57) nolu dönüőüm eőitlikleri koordinatlar için aőaęıdaki gibi yazılır;

$$\hat{\psi}_{\alpha i} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^p \frac{f_{ij}}{f_i} \hat{\phi}_{\alpha j} \quad (60)$$

$$\hat{\phi}_{\alpha i} = \frac{1}{\sqrt{\lambda_{\alpha}}} \sum_{j=1}^n \frac{f_{ij}}{f_{.j}} \hat{\psi}_{\alpha j} \quad (61)$$

Eğer $\frac{1}{\sqrt{\lambda_{\alpha}}}$ unutulursa, o zaman noktaların bir konfigürasyonunun izdüşüm noktaları ile diğer uzaya uygun konfigürasyonun izdüşüm noktalarının ağırlık merkezleri bir eksen üzerinde yer alır (38,40).

5.5. En İyi Eş Zamanlı İzdüşümün Bulunması

n satır ile p sütuna sahip (n,p)'lik kontenjans tablosu olan K çapraz tablosu ele alınsın.

Aynı eksen üzerinde satır ve sütunların bütün setleri gösterilmek istenir;

a-Her bir j sütun noktası satır noktalarının ağırlık merkezidir, her bir i satırının sahip olduğu ağırlık "i satırının parçası j sütununun profili içindedir" ve şu şekilde ifade edilir;

$$p_i = \frac{f_{ij}}{f_{.j}} \left(\sum_i p_i = 1 \right) \quad (62)$$

b-Her bir i satır noktası sütun noktalarının ağırlık merkezidir, bu sütunların her birinin sahip olduğu ağırlık "i satırlarının profilindeki j sütunlarının parçasıdır" ve şu şekilde ifade edilir;

$$p'_j = \frac{f_{ij}}{f_{.i}} \left(\sum_j p'_j = 1 \right) \quad (63)$$

Eğer ϕ_j eksen üzerindeki j sütununun absisini belirlerse (ϕ_j , bir ϕ vektörünün j'inci bileşenidir) ve eğer ψ_i aynı eksen üzerindeki i satırının absisini belirlerse o zaman aşağıdaki a ve b koşulları yazılır;

$$(a) \quad \phi = \mathbf{D}_p^{-1} \mathbf{F}' \psi \quad \text{ve} \quad \phi_j = \sum_{i=1}^n \left(\frac{f_{ij}}{f_{.j}} \right) \psi_i \quad (64)$$

$$(b) \quad \psi = \mathbf{D}_n^{-1} \mathbf{F} \phi \quad \text{ve} \quad \psi_i = \sum_{j=1}^p \left(\frac{f_{ij}}{f_{.i}} \right) \phi_j \quad (65)$$

Bu ağırlık merkezi eşitliklerini eşzamanlı olarak gerçekleştirmek genellikle mümkün değildir. Bu sebeple pozitif bir β katsayısı bulunur ve aşağıdaki eşitlik yazılır;

$$\phi = \beta \mathbf{D}_p^{-1} \mathbf{F}' \psi \quad (66)$$

$$\psi = \beta \mathbf{D}_n^{-1} \mathbf{F} \phi \quad (67)$$

β , 1'e eşit ya da daha büyüktür; diğer durumda üstteki iki eşitliğe göre; iki setin her biri eksenin bir aralığını kapsar ve bu eksen diğer set tarafından kapsanan aralıkta tam olarak içerilmiş olur. Bu yolla en küçük β bulunmuş olur.

(66) nolu eşitlik (67) nolu eşitlikte yerine yazılırsa;

$$\psi = \beta^2 \mathbf{D}_n^{-1} \mathbf{F} \mathbf{D}_p^{-1} \mathbf{F}' \psi \quad (68)$$

eşitliği elde edilmiş olur. Diğer bir deyişle ψ aşağıdaki ifadenin özvektörüdür;

$$\mathbf{D}_n^{-1} \mathbf{F} \mathbf{D}_p^{-1} \mathbf{F}'$$

En büyük özdeğeri ise ;

$$\lambda = \frac{1}{\beta^2} \quad (69)$$

olur.

Uyum analizinden elde edilmiş λ_α özdeğerleri 1'e eşit veya daha azdır.

ϕ ve ψ vektörleri, uyumdaki setlerin her birine uyan ilk faktörlerdir.

Eksen üzerinde sunulan en iyi β ağırlık merkezi için bu genelleştirilebilir ve iki ortogonal eksen ile tanımlanan bir iki boyutlu uzayda sunulan en iyi (β_1, β_2) ağırlık merkezi söz konusudur. O zaman herhangi bir boyutun (p 'den az) bir alt uzayına genişletilebilir (38).

5.6. Açıklanma Derecesi ve Kare Korelasyonların Hesaplanması

Her bir eksen için katsayıların iki serisini hesaplamak için uygulanır (bu katsayılar veri matrisinin satır ve sütunlarına eşittir).

1. Açıklanma Derecesi her bir ana eksen ile ilişkili olan her bir değişken tarafından açıklanan varyansın oranını ifade eder. Bu oran değişkenlerin bütün seti ele alınarak hesaplanır.

2. Kare korelasyonlar bir ana eksen tarafından açıklanan bir değişkenin varyansının parçasını ifade eder.

(a) Açıklanma Derecesi: u_α 'nın D_p^{-1} normu 1'e eşittir:

$$u_\alpha' D_p^{-1} u_\alpha = 1 \quad (70)$$

$u_\alpha = D_p \varphi_\alpha$ olduğundan, φ_α 'nın D_p normu 1'e eşittir:

$$\varphi_\alpha' D_p \varphi_\alpha = 1 \quad (71)$$

yani

$$\sum_{j=1}^p f_j \varphi_{\alpha j}^2 = 1 \quad (72)$$

olur. Bu durumda α ana eksenini üzerindeki \mathcal{R}^n 'nin j noktalarının görüntüsü aşağıdaki denkleme eşit olur;

$$\sum_{i=1}^n \begin{pmatrix} f_{ij} \\ f_j \end{pmatrix} \psi_{\alpha i} = \sqrt{\lambda_\alpha} \varphi_{\alpha j} = \hat{\varphi}_{\alpha j} \quad (73)$$

G'ye (ağırlık merkezi; her iki setin noktalarının ortalaması) uyumu ile birlikte, α eksenini üzerinde görüntülenmiş noktaların setinin varyansı bu nedenle aşağıdaki gibidir;

$$\sum_j f_j \hat{\varphi}_{\alpha j}^2 = \sum_j f_j (\sqrt{\lambda_\alpha} \varphi_{\alpha j})^2 = \lambda_\alpha \quad (74)$$

Bölüm (bir niceliğin birkaç kere olduğunu gösteren sayı):

$$\frac{\lambda_\alpha \varphi_{\alpha j}^2 f_j}{\lambda_\alpha} = f_j \varphi_{\alpha j}^2 = ca_\alpha(j) \quad (75)$$

α ana eksenine j elemanının tam katkısı aynı şekilde tanımlanır;

$$\sum_{j=1}^p ca_{\alpha}(j) = 1 \quad (76)$$

α ana eksenine i elemanın tam katkısı aynı şekilde tanımlanır;

$$ca_{\alpha}(i) = f_i \psi_{\alpha}^2 \quad (77)$$

ve

$$\sum_{i=1}^n ca_{\alpha}(i) = 1 \quad \text{'dir.} \quad (78)$$

(b) Kare korelasyonlar: \mathcal{R}^n 'de ağırlık merkezine j değişkeninin uzaklığı şu şekilde ifade edilir;

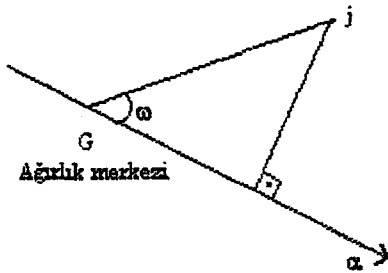
$$d^2(j, G) = \sum_{i=1}^n \frac{1}{f_i} \left(\frac{f_{ij}}{f_j} - f_i \right)^2 \quad (79)$$

α eksenindeki j değişkeninin izdüşümünün karesi aşağıdaki denkleme eşittir;

$$d_{\alpha}^2(j, G) = \left(\sqrt{\lambda_{\alpha}} \varphi_{\alpha j} \right)^2 \quad (80)$$

$$\sum_{\alpha} d_{\alpha}^2(j, G) = d^2(j, G) \quad (81)$$

olur.



Şekil-1: j 'inci eleman ile α ekseninin kare korelasyonu.

Şekil-1'de α eksenine göre j noktasının pozisyonunu gösteren nicelik;

$$\cos^2 \omega = \frac{d_{\alpha}^2(j, G)}{d^2(j, G)} = cr_{\alpha}(j) \quad (82)$$

Bu nicelik j'inci eleman ile α ekseninin kare korelasyonudur ve,

$$\sum_{\alpha} cr_{\alpha}(j) = 1 \text{ olur.} \quad (83)$$

p elemanın α eksenini üzerindeki izdüşümü;

$$\{(\sqrt{\lambda_{\alpha}} \varphi_{\alpha j}) ; j = 1, \dots, p\} \quad (84)$$

olup diğer setin n elemanına dönüştürülebilir ve böylece bu elemanların α eksenini üzerindeki izdüşümü;

$$\{(\sqrt{\lambda_{\alpha}} \psi_{\alpha i}) ; i = 1, \dots, n\} \quad (85)$$

Gerçekte α eksenini ile açıklanan değişken λ_{α} 'ya eşittir. Buradaki λ_{α} korelasyon matrisinin α 'ıncı özdeğerini verir. Fakat yeni eksen üzerindeki j'inci değişken-noktanın koordinatını $u_{\alpha j} \sqrt{\lambda_{\alpha}}$ 'ya orantılı olup aşağıdaki denklem elde edilir;

$$\sum_j (u_{\alpha j} \sqrt{\lambda_{\alpha}})^2 = \lambda_{\alpha} \quad (86)$$

Bu denklemdeki u_{α} bütündür. α ekseninin $100u_{\alpha}^2$ yüzde sayısı j değişkenini (açıklanma derecesi) ile açıklanır.

Diğer yandan \mathcal{R}^n 'deki bir j değişkeninden orijine olan uzaklık 1'e eşittir. Böylece ana eksenini ortonormalize temelinde aşağıdaki denklem elde edilir;

$$\sum_{\alpha} \lambda_{\alpha} u_{\alpha}^2 = 1 \quad (87)$$

Bu toplamdaki $\lambda_{\alpha} u_{\alpha}^2$, j değişkeninin toplam varyansının parçasını ifade eder ve bu j değişkenini α eksenini (kare korelasyon) üzerinde yer alır (38,40).

Hesaplamalar

Ağırlık merkezine olan ilişkinin analizi

Hesaplamalar \mathcal{R}^p 'de yapılmaktadır.

f_i ağırlığına sahip olan her bir i noktası şu koordinatlara sahiptir;

$$\frac{f_{ij}}{f_i} \quad j = 1, \dots, p \text{ için} \quad (88)$$

Noktalar setinin G ağırlık merkezi veya noktaların ortalaması aşağıdaki gibi hesaplanan onun j'inci bileşeni gibi hesaplanır;

$$g_j = \sum_{i=1}^n f_i \frac{f_{ij}}{f_i} = \sum_{i=1}^n f_{ij} = f_j \quad (89)$$

Analizde ağırlık merkezinin yerine aşağıdaki yeni eşitlik konularak;

$$\frac{f_{ij}}{f_i} \text{ ile } \frac{f_{ij}}{f_i} - f_j \text{ ya da } \frac{f_{ij} - f_i f_j}{f_i} \quad (90)$$

elde edilir.

Noktaların seti aşağıdaki eşitlik ile tanımlanan (p-1) boyutlu H altuzayında tanımlıdır;

$$\sum_j \left(\frac{f_{ij}}{f_i} \right) = 1 \quad \text{her } i \text{ için} \quad (91)$$

Bu altuzay G'ye uyacak şekilde analizin G ağırlık merkezini ve ana eksenini içerir.

Bu altuzaydaki her vektör bileşenlerinin toplamı sıfır olacak şekildedir. Eğer x ve y H'nin iki noktası ise ve H'nin koordinatları x_j ve y_j ise ($j=1, \dots, p$) o zaman aşağıdaki eşitlikler söz konusudur;

$$\sum_{j=1}^p x_j = \sum_{j=1}^p y_j = 1 \quad (92)$$

bu sebeple

$$\sum_{j=1}^p (x_j - y_j) = 0 \quad (93)$$

olur.

(43) nolu eşitlikteki S matrisi, S^* matrisi olur ve bu matrisin genel terimleri de s_{ij}^* olur, bunun içinde şu değişim uygulanır;

$$\frac{f_{ij}}{f_i} \text{ ile } \frac{f_{ij} - f_i f_j}{f_i}$$

böylece;

$$s_{ij}^* = \sum_{i=1}^n \frac{(f_{ij} - f_i f_j)(f_{ij'} - f_i f_{j'})}{f_i f_{j'}} \quad (94)$$

elde edilir.

s_{ij}^* değeri f_{ij} , f_i , f_j nicelikleri arasındaki ilişki ile açıklanır ve aşağıdaki denklem yazılır;

$$s_{ij}^* = \sum_{i=1}^n \left(\frac{f_{ij} f_{ij'}}{f_i f_{j'}} - f_{ij} \right) = s_{ij'} - f_j \quad (95)$$

Eğer u^* , G orijinli bir ana eksen ise bu durumda aşağıdaki denklem yazılır;

$$\sum_{j=1}^p s_{ij}^* u_j^* = \lambda^* u_j^* \quad (96)$$

ya da

$$\sum_j s_{ij}^* u_j^* - f_j \sum_j u_j^* = \lambda^* u_j^* \quad (97)$$

$u^* \in H$ olduğundan

$$\sum_j j' u_j^* = 0 \quad (98)$$

elde edilir ve bu nedenle aşağıdaki denklem elde edilir;

$$\sum_j s_{ij}^* u_j^* = \lambda^* u_j^* \quad (99)$$

H' 'ye ait olan S^{**} 'in her özdeğeri aynı özvektöre ilişkin S 'nin bir özvektörüdür. Buna ek olarak, G orijinine bağlanmış satır özdeğeri 1 olan S matrisinin bir özvektörü, ve özdeğeri 0 olan bir S^* matrisinin bir özvektörüdür.

Bu nedenle özdeğeri 1 olan S' 'yi diagonalize etmek yeterlidir.

(94) nolu eşitlikten hareketle;

$$\text{trS}^* = \sum_{i=1}^n \frac{(f_{ij} - f_i f_j)^2}{f_i f_j} \quad (100)$$

yazılır.

Toplam k olduğundan

$k \text{trS}^*$

kantitesi, \mathbf{K} matrisinin satır ve sütunlarının bağımsızlığının hipotezini test eden $(n-1)(p-1)$ serbestlik dereceli yaklaşık ki-kare dağılımı gösteren bir istatistiktir. Böylece \mathbf{C} 'nin düşük olmayan özdeğerlerinin toplamı istatistiksel bir yorum çıkarmaya sahiptir. \mathbf{S} matrisi;

$$\text{trS} = \text{trS}^* + 1 \quad (101)$$

ilişisine sahiptir (38,40).

5.7. \mathbf{S} Matrisinin Simetrikleştirilmesi

\mathbf{S} matrisi simetrik değildir. Matrisin genel terimi şöyle yazılır;

$$s_{ij} = \sum_{i=1}^n \frac{f_{ij} f_{ij'}}{f_i f_j} \quad (102)$$

Matris notasyonu ile bu matris;

$$\mathbf{S} = \mathbf{F}' \mathbf{D}_n^{-1} \mathbf{F} \mathbf{D}_p^{-1} \quad (103)$$

şeklindedir.

$$\hat{\mathbf{A}} = \mathbf{F}' \mathbf{D}_n^{-1} \mathbf{F} \quad (104)$$

matrisi simetriktir ve \mathbf{D}_p^{-1} matrisi diagonaldir. Böylece aşağıdaki eşitlikler yazılabilir;

$$\frac{1}{f_j} = \left(\frac{1}{\sqrt{f_j}} \right) \left(\frac{1}{\sqrt{f_j}} \right) \quad (105)$$

veya

$$\mathbf{D}_p^{-1} = \mathbf{D}_p^{-1/2} \mathbf{D}_p^{-1/2} \quad (106)$$

Buradan;

$$\mathbf{S} = \hat{\mathbf{A}} \mathbf{D}_p^{-1/2} \mathbf{D}_p^{-1/2} \quad (107)$$

(43) nolu eşitlikten hareketle;

$$\hat{\mathbf{A}} \mathbf{D}_p^{-1/2} \mathbf{D}_p^{-1/2} \mathbf{u} = \lambda \mathbf{u} \quad (108)$$

elde edilir.

(108) nolu eşitliğin her iki tarafı $\mathbf{D}_p^{-1/2}$ ile çarpılıp;

$$\mathbf{D}_p^{-1/2} \mathbf{u} = \omega \quad (109)$$

kabul edilirse bu durumda aşağıdaki denklem elde edilmiş olur;

$$\mathbf{D}_p^{-1/2} \hat{\mathbf{A}} \mathbf{D}_p^{-1/2} \mathbf{w} = \lambda \mathbf{w} \quad (110)$$

Böylece;

$$\mathbf{S}_1 = \mathbf{D}_p^{-1/2} \hat{\mathbf{A}} \mathbf{D}_p^{-1/2} \quad (111)$$

matrisi simetriktir ve \mathbf{S} matrisi gibi aynı λ öz değerine sahiptir. Aşağıdaki eşitlikten hareketle ϕ faktörünü yazmak mümkündür.

$$\mathbf{u} = \mathbf{D}_p^{1/2} \mathbf{w} \quad (112)$$

ϕ faktörü ;

$$\phi = \mathbf{D}_p^{-1} \mathbf{u} = \mathbf{D}_p^{-1/2} \mathbf{w} \quad (113)$$

şeklinde elde edilir (38,40).

NOT: \mathcal{R}^p 'deki ilk koordinatlar üzerinde aşağıdaki transformasyon uygulanır;

$$\mathbf{y} = \mathbf{D}_p^{-1/2} \mathbf{x} \quad (114)$$

x_1 ve x_2 gibi iki nokta arasındaki ki-kare uzaklığı yeni koordinatlar ile ;

$$d^2(x_1, x_2) = (x_1 - x_2)' D_p^{-1} (x_1 - x_2) \quad (115)$$

şeklinde olur ve genel uzaklık ;

$$d^2(y_1, y_2) = (y_1 - y_2)' (y_1 - y_2) \quad (116)$$

olur. Simetrik S_1 matrisinin direkt diagonalizasyonunu gösterir (38,40).

6. HOMOJENİTE ANALİZİ

Kategorik çok değişkenli verileri ölçmek için çok çeşitli metotlar söz konusudur. Bu metotlar farklı ülkelerde, farklı yazarlarca, farklı adlar altında sunulur. Nesnelere (gruplara) ve kategorilere nümerik değerler atama yoluyla nominal (kategorik) verileri analiz eden HA farklı yazarlarca değişik adlarla anılmaktadır. Amerika'da Optimal Scaling, Optimal Scoring, Appropriate Scoring; Kanada'da Dual Scaling, Dutch Homogeneity Analysis; İsrail'de Scalogram Analysis; Fransa'da ise Multiple Correspondence Analysis ve Japonya'da ise Quantification Method olarak adlandırılır. HA, çoklu kontenjans Tablosu şeklindeki ikiden daha fazla kategorik yapılu değişkenlerin analizinde uygun bir yöntemdir (16,17,18,40,41).

Uyum analizi iki yönlü çapraz tabloların analizi için kullanılırken, ikiden daha fazla değişken söz konusu olduğu durumlarda çoklu çapraz tabloların analizinde HA kullanılır. HA, CA'nın genişletilmiş şekli olduğundan, iki yönlü çapraz tabloların analizinde CA ile aynı sonuçları vermektedir. Böylece, HA ile iki ya da daha çok değişken arasındaki ilişkiyi analiz etmek ve grafiksel gösterim ile özetlemek mümkündür.

HA, değişkenlerin kategorilerini ya da grupları noktalar halinde grafiksel gösterimlerle ifade eder. Kolonla ilgili olarak kategori sayısı kadar skorlar oluşturulur. Tablodaki veri çiftlerinden hesaplanmış korelasyonları maksimize etmek amacıyla birimler için de gözlem sayısı kadar uygun skorlar oluşturulur. Kolonlar olarak tanımlanan değişkenlerin kategorileri ve satırlar olarak tanımlanan gruplarla ayırıcı nitelikte olan matrisler oluşturulur. Ayırıcı matris (ikili giriş matris) açıklamaları ve yöntemin kolaylaştırılmasını sağlar. HA ile aynı kategori çevresinde hangi nesnelere birbirine yakın, hangi nesnelere ise uzak olduğunu grafiksel olarak ifade eder. Analizde bütün

değişkenler için bu yapılır. Bu yolla kategoriler nesnelere homojen olarak alt gruplara ayırır (4).

Değişken sayısı Q olarak verilsin. Her bir değişken q ve her bir değişkenin aldığı değer (cevap) p_q olsun.

Cevap kategorilerinin toplam sayısı p aşağıdaki gibi ifade edilir;

$$p = \sum_{q=1}^Q p_q \quad (117)$$

Q değişkenli birimlerin sayısı n 'dir.

İkili kodlamayla $(0,1)$ n birimin değerleri p sütun ve n satırlı ile Z matrisiyle ifade edilir. Z matrisi Q adet alt matrisin yan yana getirilmiş şeklidir ve matris notasyonu ile aşağıdaki gibi ifade edilir:

$$Z = [Z_1, Z_2, \dots, Z_q, \dots, Z_Q] \quad (118)$$

Z_q alt matrisleri n satır ve p_q kolondan oluşup, her bir değişken için verilen cevapların dummy $(0,1)$ olarak verildiği matrislerdir.

(n,Q) boyutlu R matrisi, değişkenlerin değerlerinin sayısal olarak kodlandığı $(1,2,3,..)$ matris olup Z matrisinin özetidir (38).

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1Q} \\ r_{21} & r_{22} & \dots & r_{2Q} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{nQ} \end{bmatrix}_{n \times Q} \quad i=1,2,\dots,n \quad q=1,2,\dots,Q \quad (119)$$

6.1. Z ile bağlantılı Burt tablosu

$Z = [Z_1, Z_2, \dots, Z_Q]$ olan Z matrisiyle ilgilenilirse, kare matris

$$B = Z'Z \quad (120)$$

olur ve Z ile ilişkili Burt'ün kontenjans tablosu olarak isimlendirilir.

B Matrisi Q^2 bloklarından oluşur.

q 'ncu $Z'_q Z_q$ kare matris (p_q, p_q) boyutlu bir diagonal matristir.

(q, q') indeksli $Z'_q Z_q$ bloğu, q ve q' olan iki değişkenin değerlerinin çapraz tablosudur.

D diagonal matrisi B gibi aynı diagonal elemanlara sahip olan (p, p) boyutlu bir matristir. Bu matrisin diagonal elemanları kategorilerin her birinin frekanslarını içerir.

$$D_q = Z'_q Z_q \quad (121)$$

diagonal matrisinin elemanları q değişkeninin çeşitli kategorilerinin frekanslarına eşittir (38).

6.2. İki Değişkenli Uyum

Cevap matrisi

$$Z = [Z_1 | Z_2] \quad (122)$$

olarak yazılır ve kategoriler arasındaki ilişkileri açıklamada aşağıdakileri hesaplamak için kullanılır (38).

1. (n, p) boyutlu Z matrisinin uyum analizi
2. (p, p) boyutlu B matrisinin uyum analizi
3. (p_1, p_2) boyutlu $Z'_1 Z_2$ matrisinin uyum analizi
4. Z_1 ve Z_2 iki sütun bloklarının kanonik analizi

6.3. Birinci Eşitlik (1 ve 2 arasındaki)

α 'ncı faktör (ϕ_α)

$$\frac{1}{Q} D^{-1} Z' Z \phi_\alpha = \mu_\alpha \phi_\alpha \quad (123)$$

olup, CA için kabul edilen notasyon, (123) nolu eşitliğin sol kısmındaki matris

$$D_p^{-1} F' D_n^{-1} F \quad (124)$$

olarak yazılır,

ve burada;

$$\mathbf{F} = \frac{1}{n\mathbf{Q}} \mathbf{Z} \quad (125)$$

$$\mathbf{D}_p = \frac{1}{n\mathbf{Q}} \mathbf{D} \quad (126)$$

$$\mathbf{D}_n = \frac{1}{n} \mathbf{I}_n \quad (127)$$

(\mathbf{I}_n (n,n) boyutlu birim matristir)

$$\mathbf{D}_p^{-1} \mathbf{F}' \mathbf{D}_n^{-1} \mathbf{F} \phi_\alpha = \mu_\alpha \phi_\alpha \quad (128)$$

bu eşitlik (123) nolu eşitlik dikkate alınarak yazılır.

\mathbf{B} matrisi (120) nolu eşitlikteki gibi alındığında \mathbf{B} simetriktir ve marjinal satır ve sütunları \mathbf{QD} matrisinin diagonal elemanlarıdır.

\mathbf{B} analiz edildiğinde yeni bir \mathbf{F} matrisi bulunur;

$$\mathbf{F} = \frac{1}{n\mathbf{Q}^2} \mathbf{B} \quad (129)$$

ve ilgili yeni matrisler \mathbf{D}_n ve \mathbf{D}_p ;

$$\mathbf{D}_n = \mathbf{D}_p = \frac{1}{n\mathbf{Q}} \mathbf{D} \text{ 'ye eşittir.} \quad (130)$$

Bu nedenle diagonalize edilen matris bu durumda aşağıdaki gibi yazılır;

$$\frac{1}{\mathbf{Q}^2} \mathbf{D}^{-1} \mathbf{B}' \mathbf{D}^{-1} \mathbf{B} \quad (131)$$

(123) nolu eşitliğin her iki tarafı $\left(\frac{1}{\mathbf{Q}}\right) \mathbf{D}^{-1} \mathbf{B}$ ile çarpıldığında

$$\frac{1}{\mathbf{Q}^2} \mathbf{D}^{-1} \mathbf{B} \mathbf{D}^{-1} \mathbf{B} \phi_\alpha = \mu_\alpha^2 \phi_\alpha \quad (132)$$

elde edilir (38).

6.4. İkinci Eşitlik (1 ve 3 arasındaki)

$(\phi_\alpha, \psi_\alpha)$ faktörlerinin her bir çifti için, $Z_1'Z_2$ kontenjans tablosunun analizden elde edilen aynı λ_α özdeğeriyle ilişkili olduğu gösterilsin. Öyle ki Z ya da B 'nin analiz edilmesinden bir ilişki faktörü ϕ_α vardır.

$$\phi_\alpha = \begin{bmatrix} \phi_n \\ \psi_n \end{bmatrix} \quad (133)$$

$$D_1 = Z_1'Z_1 \quad (134)$$

ve

$$D_2 = Z_2'Z_2 \quad (135)$$

eşitliğinden;

$$D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} \quad (136)$$

matrisi yazılabilir. Burada D_1 ve D_2 'nin diagonal elementleri $Z_1'Z_2$ matrisinin marjinal satır ve sütunlarıdır.

Bu tablonun analizi, çift geçişli eşitliğe götürür;

$$\phi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} D_1^{-1} Z_1' Z_2 \psi_\alpha \quad (137)$$

$$\psi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} D_2^{-1} Z_2' Z_1 \phi_\alpha \quad (138)$$

Bu eşitlikler eşitliklerin bir sistemi olarak yazılır;

$$D_1^{-1} (D_1 \phi_\alpha + Z_1' Z_2 \psi_\alpha) = (1 + \sqrt{\lambda_\alpha}) \phi_\alpha \quad (139)$$

$$D_2^{-1} (D_2 \psi_\alpha + Z_2' Z_1 \phi_\alpha) = (1 + \sqrt{\lambda_\alpha}) \psi_\alpha \quad (140)$$

$$\begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}^{-1} \begin{bmatrix} D_1 & Z_1' Z_2 \\ Z_2' Z_1 & D_2 \end{bmatrix} \begin{bmatrix} \phi_\alpha \\ \psi_\alpha \end{bmatrix} = (1 + \sqrt{\lambda_\alpha}) \begin{bmatrix} \phi_\alpha \\ \psi_\alpha \end{bmatrix} \quad (141)$$

(141) nolu eşitliğin iki tarafı $1/Q$ ($Q=2$) ile çarpıldığında;

$$\frac{1}{Q} \mathbf{D}^{-1} \mathbf{Z}' \mathbf{Z} \phi_{\alpha} = \left(\frac{1 + \sqrt{\lambda_{\alpha}}}{2} \right) \phi_{\alpha} \quad (142)$$

eşitliği elde edilir.

$$\mu_{\alpha} = \frac{1 + \sqrt{\lambda_{\alpha}}}{2} \quad (143)$$

ile eşitlik

$$\frac{1}{Q} \mathbf{D}^{-1} \mathbf{Z}' \mathbf{Z} \phi_{\alpha} = \mu_{\alpha} \phi_{\alpha} \quad (144)$$

şeklinde ifade edilir.

Eğer λ_{α} , $\mathbf{Z}'_1 \mathbf{Z}_2$ 'nin analizinden elde edilen α 'ncı en büyük özdeğer ise o zaman (143) nolu eşitlik \mathbf{Z} 'nin analizinin en büyük özdeğerini verir (38,40).

6.5. İkiden Daha Fazla Değişken İçin Genelleme

\mathcal{R}^n 'nin p noktalarına karşılık gelen $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_q, \dots, \mathbf{Z}_Q]$ matrisi p kolona sahiptir. Her bir alt matris \mathbf{Z}_q , p_q boyutlu lineer bir γ_q altuzayı meydana getirir.

Bütün bu lineer altuzaylar en azından ortaklaşa birinci açıortaya sahiptirler. \mathbf{Z} matrisinin rankı bu yüzden ;

$$p - (Q-1) \text{ 'e eşittir.} \quad (145)$$

p_q bileşenlerinin vektörü olan \mathbf{Q}_q , \mathbf{Z}_q 'nun kolonlarıyla tanımlanmış, temelde γ_q 'nun bir \mathbf{m}_q noktasının koordinatlarıdır.

\mathcal{R}^n 'deki \mathbf{m}_q 'nun koordinatları

$$\mathbf{m}_q = \mathbf{Z}_q \phi_q \quad (146)$$

eşitliğinin bileşenleridir.

Bu nokta \mathbf{m}_q 'nun uzaklığının karesi ya da onun öklid uzaklığı şu şekilde verilebilir;

$$\phi_q' \mathbf{Z}_q' \mathbf{Z}_q \phi_q = \phi_q' \mathbf{D}_q \phi_q \quad (147)$$

İki sorulu q ve q' çapraz tablodaki kontenjans tablosunun CA γ_q ve $\gamma_{q'}$ altuzaylarının ilişkili durumlarını araştırmak için indirgenir.

İkili geçiş eşitlikleri (137) nolu ve (138) nolu eşitlikler aşağıdaki gibi yazılır (notasyonu basitleştirmek için sıra ile α indeksi ihmal edilir) ;

$$\Phi_q = \frac{1}{\sqrt{\lambda}} \mathbf{D}_q^{-1} \mathbf{Z}'_q \mathbf{Z}_{q'} \Phi_{q'} \quad (148)$$

$$\Phi_{q'} = \frac{1}{\sqrt{\lambda}} \mathbf{D}_{q'}^{-1} \mathbf{Z}'_{q'} \mathbf{Z}_q \Phi_q \quad (149)$$

Bu eşitliklerden takip eden eşitlikler çıkarılır;

$$\mathbf{Z}_q \Phi_q = \frac{1}{\sqrt{\lambda}} \mathbf{Z}_q \mathbf{D}_q^{-1} \mathbf{Z}'_q \mathbf{Z}_{q'} \Phi_{q'} \quad (150)$$

$$\mathbf{Z}_{q'} \Phi_{q'} = \frac{1}{\sqrt{\lambda}} \mathbf{Z}_{q'} \mathbf{D}_{q'}^{-1} \mathbf{Z}'_{q'} \mathbf{Z}_q \Phi_q \quad (151)$$

ve

$$\mathbf{m}_q = \frac{1}{\sqrt{\lambda}} \mathbf{P}_q \mathbf{m}_{q'} \quad (152)$$

$$\mathbf{m}_{q'} = \frac{1}{\sqrt{\lambda}} \mathbf{P}_{q'} \mathbf{m}_q \quad (153)$$

burada;

$$\mathbf{P}_q = \mathbf{Z}_q (\mathbf{Z}'_q \mathbf{Z}_q)^{-1} \mathbf{Z}'_q \quad (154)$$

$$\mathbf{P}_{q'} = \mathbf{Z}_{q'} (\mathbf{Z}'_{q'} \mathbf{Z}_{q'})^{-1} \mathbf{Z}'_{q'} \quad (155)$$

olup, \mathbf{P}_q ve $\mathbf{P}_{q'}$ matrisleri γ_q ve $\gamma_{q'}$ 'deki izdüşüm operatörlerini gösterir.

(152) nolu ve (153) nolu eşitliklerdeki; $\gamma_{q'}$ 'deki ortogonal izdüşümündeki durum $\mathbf{m}_{q'}$ ile aynı çizgi üzerindedir. (Ve aynı şekilde γ_q 'deki \mathbf{m}_q için de durum aynıdır). Buna rağmen, $[\mathbf{Z}_q]$ ve $[\mathbf{Z}_{q'}]$ 'nün kanonik analizi aşağıdaki formülle hesaplanır.

\mathbf{m}_q ve $\mathbf{m}_{q'}$ gibi iki noktanın orijine olan kare uzaklıklarının toplamının ortalaması;

$$\varphi'_q \mathbf{D}_q \varphi_q + \varphi'_{q'} \mathbf{D}_{q'} \varphi_{q'} = 2n \quad (156)$$

olup $\mathbf{m} = \mathbf{m}_q + \mathbf{m}_{q'}$ noktasının orijine olan uzaklığı maksimumdur.

Bu uzaklığın karesi ;

$$\|\mathbf{m}^2\| = \varphi'_q \mathbf{D}_q \varphi_q + \varphi'_{q'} \mathbf{D}_{q'} \varphi_{q'} + 2\varphi'_q \mathbf{Z}'_q \mathbf{Z}_{q'} \varphi_{q'} \quad (157)$$

Maksimum yapılan $\|\mathbf{m}^2\|$, (156) nolu eşitlikten;

$$\varphi'_q \mathbf{D}_q \varphi_q = \varphi'_{q'} \mathbf{D}_{q'} \varphi_{q'} = n \quad (158)$$

elde edilir.

Bu eşitliklerden hareketle iki değişkenden daha fazla değişkene genelleme kolaylıkla yapılabilir.

$\varphi_1, \varphi_2, \dots, \varphi_Q$ olarak ifade edilen Q noktanın bileşenlerinin vektörleri $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_q; \mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_q$ 'ya temeldeki

$$\mathbf{m} = \mathbf{m}_1 + \mathbf{m}_2 + \dots + \mathbf{m}_Q \quad (159)$$

olarak alınırsa;

$$\|\mathbf{m}^2\| = \sum \left\{ \left(\varphi'_q \mathbf{Z}'_q \mathbf{Z}_{q'} \varphi_{q'} \mid q \in Q; q' \in Q \right) \right\} \quad (160)$$

eşitliği ile;

$$\sum \left\{ \left(\varphi'_q \mathbf{D}_q \varphi_q \mid q \in Q \right) \right\} = Qn \quad (161)$$

olur. p bileşenli ϕ vektörü $\phi' = \{\phi'_1, \phi'_2, \dots, \phi'_Q\}$ olarak tanımlanır.

$$\phi' \mathbf{B} \phi \quad (162)$$

(162) nolu eşitlik;

$$\phi' \mathbf{D} \phi = Qn \quad (163)$$

biçiminde ifade edilir. Böylece ϕ faktörleri en büyük özdeğerli $D^{-1}B$ 'nin özvektörleridir (38).

6.6. Çoklu Analizlerin Özellikleri

Z matrisinin analizinden elde edilen ϕ faktörleri hatırlanacak olursa;

$$\frac{1}{Q} D^{-1} B \phi = \mu \phi \quad (164)$$

q soruyla bağlantılı ϕ 'nın ϕ_q bileşenlerini göstermek için bu eşitliğin terimlerini tekrar düzenleme yoluyla; D ve B matrislerinin bloklarıyla birlikte

$$\frac{1}{Q} \sum \{ D_q^{-1} Z_q' Z_q \phi_q | q \in Q \} = \mu \phi_q \quad (165)$$

eşitliği elde edilir (38).

6.7. ϕ_q Bileşenlerinin Merkezinin Belirlenmesi

Bir q değişkeninin p_q kategorileriyle ilgili noktaların Q alt setleri aynı ağırlık merkezine sahiptir. J_q , q değişkeni ile ilgili j indeksinin p değerlerinin alt setlerini ifade eder. (J_q , p_q unsurlarına sahiptir).

q değişkeni ile bağlantılı noktaların alt setlerinin koordinatları $Z_q D_q^{-1}$ 'nin kolonlarıdır ve $\left(\frac{1}{n}\right) D_q$ 'nin diagonal unsurları alt set q 'nin p_q noktalarının bağlı parçalarıdır.

G_g ağırlık merkezinin i 'inci koordinatı;

$$g_{gi} = \sum_{j \in J_q} \frac{d_{ij}}{n} \frac{z_{ij}}{d_{ij}} = \frac{1}{n} \quad \left(\sum_{j \in J_q} z_{ij} = 1 \right) \quad (166)$$

Böylece g_{gi} ; $q: g_{gi} = g_i$ 'nin bağımsızdır.

Önemsiz olmayan faktörlerle ilgili ϕ_q bileşenleri, G 'deki orijinin dönüşümünden sonra noktaların setinin analiziyle ilgili bu faktörlerden dolayı bir merkezde toplanır (38,40).

6.8. Bir Kategori ve Bir Değişkene Uygun Varyansın Oranı

Eşitlik (165)'de verilen toplam varyans, önemsiz olmayan özdeğerlerin toplamı olup;

$$\frac{P}{Q} - 1 \quad \text{'e eşittir.} \quad (167)$$

Genellikle; bütün değişkenler iki cevap kategorisine ($p=2Q$) sahip olduğu zaman yukarıdaki eşitlik 1'e eşittir.

Bir j kategori noktasının \mathcal{R}^n 'deki ağırlık merkezinden olan uzaklığının karesi;

$$d^2(G, j) = n \sum_{i=1}^n \left(\frac{z_{ij}}{d_{ij}} - \frac{1}{n} \right)^2 \quad (168)$$

olarak yazılır ve;

$$\sum_{i=1}^n z_{ij} = d_{ij} \quad (169)$$

eşitliği hesap içine alınarak

$$d^2(G, j) = n \left(\frac{1}{d_{ij}} - \frac{1}{n} \right) \quad (170)$$

elde edilir.

j kategorisinin toplam varyansa katkısı;

$$c(j) = \frac{d_{ij}}{nQ} d^2(G, j) = \frac{1}{Q} \left(1 - \frac{d_{ij}}{n} \right) \quad (171)$$

q değişkeninin toplam varyansa katkısı aşağıdaki gibidir.

$$C(q) = \sum \left\{ c(j) \mid j \in J_q \right\} = \frac{1}{Q} (p_q - 1) \quad (172)$$

Bu nedenle bir değişkene ilişkin varyansın oranı cevap kategorilerinin sayısının artan bir fonksiyonudur. En küçük oran ($1/Q$); değişken sadece iki kategoriye sahip olduğu durumlarda elde edilir (38,40).

$$\sum_{q=1}^Q C(q) = \frac{P}{Q} - 1 \quad (\text{toplam varyans}) \quad (173)$$

6.9. \mathfrak{R}^n 'deki p Kategorinin Konfigürasyonunun Boyutlandırılması

\mathfrak{R}^n 'deki kategorilerin koordinatları \mathbf{ZD}^{-1} 'in kolonlarıdır. Onlar, \mathbf{ZD}^{-1} 'in rank boyutunun bir alt setini meydana getirir. Bu yüzden;

$$\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_Q] \quad (174)$$

şeklinde yazılır.

γ_q alt uzaylarının tümü \mathbf{Z}_q matrislerinin kolonlarıyla genelleştirilmiş olarak yaygın olarak birinci açıortaya Δ sahiptir. Bu yüzden \mathbf{Z} 'nin maksimum derecesi ;

$$p_1 + (p_2 - 1) + \dots + (p_Q - 1) = p - Q + 1 \quad (175)$$

olur.

Bu nedenle $\mathbf{D}^{-1}\mathbf{Z}'\mathbf{Z}$ matrisinin maksimum rankı $p-Q+1$ olarak diagonalize edilir. Fakat \mathbf{O} orijinli birinci açıortay Δ ; 1 özdeğeriyle ilgili bir özvektördür. Analizde, \mathbf{G} ağırlık merkezinde $(p-Q)$ tane sıfır olmayan özdeğerler bulunur.

Böylece nokta setlerinin alt uzayında bir $(p-Q)$ matrisi için özvektörler ve özdeğerleri bulabilmek için indirgeme yapılır (38,40).

6.10. En İyi Eşzamanlı Gösterim

Aynı eksen üzerindeki p kategori ve n birimin absisleri araştırılıyorsa, \mathbf{Z} 'nin analiziyle ilişkili ikili dönüşüm eşitlikleri aşağıdaki gibi verilir;

$$\hat{\psi} = \frac{1}{\sqrt{\mu}} \frac{1}{Q} \mathbf{Z} \hat{\phi} \quad (176)$$

$$\hat{\phi} = \frac{1}{\sqrt{\mu}} \mathbf{D}^{-1} \mathbf{Z}' \hat{\psi} \quad (177)$$

burada $\hat{\psi}_i$, i bireyin absisi; $\hat{\phi}_j$, j kategorinin absisidir (38,40).

7. GRAFİKSEL GÖSTERİM

HA'nın önemli özelliklerinden birisi de çoklu kontenjans tablosundaki satır ve sütunların birbirleri ile olan bağımlılıklarını grafiksel olarak ifade etmesidir. Satır ve sütun sayısının fazla olduğu kontenjans tablolarında satır ve sütunları yüzdeler ile ya da ki-kare analizleri ile karşılaştırmak zor hatta imkansız olduğundan grafiksel gösterim satır ve sütunları karşılaştırmayı kolaylaştırmaktadır. Grafikte satır ve sütunlar arasındaki bağımlılıkları daha net olarak görmek mümkündür.

Daha önceden belirlenen bir yüzde varyans açıklayıcılığına (%90, %95, %99) göre, özdeğerlerin özdeğerler toplamına olan oranının yığılımlı toplamı, belirlenen yüzde varyans açıklayıcılığına eşit ya da daha fazla olduğu özdeğerlerin sayısı, çizilecek olan grafiğin eksen sayısını verir (8,16,31,34,35,38,39,40,41,42,44).

Bu analiz yöntemi noktaların bir setinin elemanları arasındaki uzaklıkları yorumlamada iyi bir yöntem olup, eksenin ortasında yerleşen ağırlık merkezi G, her iki setin noktalarının ortalamalarını ifade eder.

GEREÇ VE YÖNTEM

1. VERİLER

Bu çalışmada 1996 yılında Denizli il merkezinde yaşayan 79211 aile araştırma evrenini oluşturmaktadır. Toplumumuzda akraba evliliği yapma oranı $p=0.116$ olarak alınmış ve kesinlik kararı $H=0.02$, güven olasılığı 0.95 ($z=1.96$) alınarak $n>985$ olarak belirlenmiştir. Örnek olarak alınacak mahallelerin seçiminde tabakalı örnekleme ve anket uygulanacak ailelerin seçiminde ise kümeleme örnekleme yöntemi kullanılmıştır. Araştırmamızda akraba evliliği yapmaya ve çocuğun özürlü olmasına neden olan etmenlerin belirlenmesi için 1000 aileye anket uygulanmış ve akraba evliliğini ve özürlü çocuğa sahip olmayı etkileyebilecek olan "Annenin Eğitim Durumu (AED)", "Annenin Doğum Yeri (ADY)", "Babanın Eğitim Durumu (BED)", "Babanın Doğum Yeri (BDY)", "Eşler Arasında Akraba Evliliği (EAA)" ve "Çocuklarında Özür Olup-olmadığı (COO)" değişkenleri ele alınmıştır.

2. ANALİZ YÖNTEMİ

Bu çalışmada geriye doğru eleminasyon yöntemi (geriye doğru aşamalı LLA), temel modelden uygun model seçimi prosedürü olarak kullanılmıştır. Bu yöntemde her adımda, şu anda kullanılmakta olan ve ondan sadece bir etki farklı olan tüm mümkün hiyerarşik modelleri kapsar. Bu prosedürde LR ki-kare içindeki en az anlamlı değişim ile ilgili etki modelden çıkartılır (6).

Bu prosedür, hücre frekanslarının en az sayıda değişkeni içeren modelini bulmak için adım adım işlem uygular. Bu prosedür, doyurulmuş modellerle başlar ve en yüksek derece etkilerini test eder. Şayet bir etki kaldırılırsa, etki geçerli olmaksızın kümelenmiş model geçerlidir ve geri kalan etkilerin herhangi birisinin kaldırılıp kaldırılamayacağını görmek için test eder. Bu işlem artık kaldırılacak etki kalmayınca kadar devam eder(4).

Bu çalışmada da geriye doğru eleminasyon yöntemi kullanılmış ve bunun sonucu olarak en uygun model bulunmuştur. Bu prosedürde iterasyon sayısı mümkün olduğunca yüksek tutulmuştur. Çünkü daha objektif sonuçlara ulaşmak için bu gerekmektedir. Geriye doğru eleminasyon yöntemi ile en uygun model bulunduktan sonra en uygun modele HA uygulanmış ve daha az işlem karmaşası ile, daha net ve kolay anlaşılır sonuçlara ulaşılmıştır.

BULGULAR

1. GERİYE DOĞRU AŞAMALI LOG-LİNEER ANALİZİN SONUÇLARI

Denizli ilinde akraba evliliği ve özürlü çocuğa sahip olma ile ilgili bir araştırmanın "Annenin Eğitim Durumu", "Annenin Doğum Yeri", "Babanın Eğitim Durumu", "Babanın Doğum Yeri", "Eşler Arasında Akraba Evliliği" ve "Çocuklarında Özür Olup-olmadığı" değişkenlerine geriye doğru aşamalı LLA uygulanmış ve en fazla 3'lü etkileşimlerin istatistiksel olarak anlamlı olduğu bulunmuştur (Tablo-3).

Tablo-3: k yönlü etkilerin LR Ki-kare istatistik sonuçları.

K	DF	L.R. Kikare	P	Pearson Chisq	P
1	14	2583,420	,0000	20648,612	,0000
2	77	1703,506	,0000	3958,177	,0000
3	212	253,898	,0258	447,595	,0000
4	308	75,084	1,0000	65,301	1,0000
5	224	1,941	1,0000	1,030	1,0000
6	64	,000	1,0000	,000	1,0000

En uygun model bulunmadan önce AED, ADY, BED, BDY, EAA ve COO değişkenlerine ilişkin doymuş modelin mümkün bütün etkileşimlerinin LR- χ^2 değerleri ve önemlilik düzeyleri Tablo-4'de verilmiştir.

Tablo-4: En uygun model bulunmadan önce ilgili etkileşimlere ait serbestlik dereceleri, kısmi ki-kare değerleri ve önemlilik düzeyleri.

Etki Adı	SD	LR- χ^2	p
ADY*AED*BDY*BED*EAA	64	1,999	1,0000
ADY*AED*BDY*BED*COO	64	,000	1,0000
ADY*AED*BDY*EAA*COO	16	,000	1,0000
ADY*AED*BED*EAA*COO	32	,000	1,0000
ADY*BDY*BED*EAA*COO	16	,000	1,0000
AED*BDY*BED*EAA*COO	32	,000	1,0000
ADY*AED*BDY*BED	64	22,925	1,0000
ADY*AED*BDY*EAA	16	,375	1,0000
ADY*AED*BED*EAA	32	5,798	1,0000
ADY*BDY*BED*EAA	16	9,413	,8954
AED*BDY*BED*EAA	32	11,226	,9997
ADY*AED*BDY*COO	16	,020	1,0000
ADY*AED*BED*COO	32	,026	1,0000
ADY*BDY*BED*COO	16	,123	1,0000
AED*BDY*BED*COO	32	,008	1,0000
ADY*AED*EAA*COO	8	,168	1,0000
ADY*BDY*EAA*COO	4	,167	,9967
AED*BDY*EAA*COO	8	,185	1,0000

ADY*BED*EAA*COO	8	.000	1.0000
AED*BED*EAA*COO	16	.000	1.0000
BDY*BED*EAA*COO	8	.126	1.0000
ADY*AED*BDY	16	29,711	.0196
ADY*AED*BED	32	25,819	.7715
ADY*BDY*BED	16	45,848	.0001
AED*BDY*BED	32	25,736	.7752
ADY*AED*EAA	8	25,886	.0011
ADY*BDY*EAA	4	22,503	.0002
AED*BDY*EAA	8	20,370	.0090
ADY*BED*EAA	8	12,285	.1389
AED*BED*EAA	16	10,449	.8422
BDY*BED*EAA	8	10,480	.2329
ADY*AED*COO	8	7,411	.4930
ADY*BDY*COO	4	8,096	.0881
AED*BDY*COO	8	8,634	.3741
ADY*BED*COO	8	5,047	.7525
AED*BED*COO	16	3,862	.9991
BDY*BED*COO	8	7,882	.4451
ADY*EAA*COO	2	9,831	.0073
AED*EAA*COO	4	4,287	.3685
BDY*EAA*COO	2	3,548	.1697
BED*EAA*COO	4	7,898	.0954
ADY*AED	8	35,697	.0000
ADY*BDY	4	616,185	.0000
AED*BDY	8	11,509	.1745
ADY*BED	8	11,169	.1923
AED*BED	16	611,175	.0000
BDY*BED	8	19,295	.0134
ADY*EAA	2	4,572	.1017
AED*EAA	4	15,276	.0042
BDY*EAA	2	3,745	.1537
BED*EAA	4	2,486	.6471
ADY*COO	2	.576	.7497
AED*COO	4	6,357	.1740
BDY*COO	2	5,574	.0616
BED*COO	4	.855	.9310
EAA*COO	1	13,803	.0002
ADY	2	61,589	.0000
AED	4	566,642	.0000
BDY	2	120,135	.0000
BED	4	452,235	.0000
EAA	1	428,587	.0000
COO	1	954,234	.0000

Tablo-4 incelendiğinde ADY*AED*BDY, ADY*BDY*BED, ADY*BDY*EAA, ADY*AED*EAA, AED*BDY*EAA, ADY*EAA*COO, ADY*AED, ADY*BDY, AED*BED, BDY*BED, AED*EAA, EAA*COO etkileşimlerinin ve ADY, AED, BDY, BED, EAA, COO ana etkilerinin istatistiksel olarak önemli olduğu bulunmuştur.

Geriye doğru aşamalı LLA sonucunda bulunan en uygun model;

$$\ln(f_{ijklmn}) = \mu + \alpha_{ACE} + \alpha_{ADF} + \alpha_{DCE} + \alpha_{BF} + \alpha_{AB} + \alpha_{BC}$$

modeli olup, burada;

A:ADY:Annenin Doğum Yeri (Köy, İlçe, İl)

B:AED:Annenin Eğitim Durumu (Okuryazar Değil+Okuryazar, İlk, Orta, Lise, Yüksekokul)

C:EAA:Eşler Arasında Akraba Evliliği (Var, Yok)

D:BDY:Babanın Doğum Yeri (Köy, İlçe, İl)

E:COO:Çocukta Özür Olup Olmadığı (Var, Yok)

F:BED:Babanın Eğitim Durumu (Okuryazar Değil+Okuryazar, İlk, Orta, Lise, Yüksekokul)

olarak alınmıştır.

Geriye doğru aşamalı LLA sonucunda bulunan en uygun model ve bu modelin etkileşimlerinin LR- χ^2 sonuçları Tablo-5'te verilmiştir.

Tablo-5: En uygun modelin etkileşimlerinin LR ki-kare sonuçları.

Kısmi İlişki	SD	L.R.Kikare	p
ADY*EAA*COO	2	6,358	,0416
ADY*BDY*BED	16	56,348	,0000
AED*BED	16	639,364	,0000
BDY*EAA*COO	2	6,712	,0349
ADY*AED	8	56,981	,0000
AED*EAA	4	41,110	,0000

Geriye doğru aşamalı LLA sonucunda elde edilen en uygun modelin ADY*EAA*COO, ADY*BDY*BED, BDY*EAA*COO, AED*BED, ADY*AED ve AED*EAA etkileşimleri istatistiksel olarak anlamlı bulunmuştur. En uygun model bulunmadan önce istatistiksel olarak anlamlı bulunan etkileşimler en uygun model bulunduktan sonra yine anlamlı olarak bulunmasına karşılık, en uygun model bulunmadan önce istatistiksel olarak anlamlı bulunmamış olan BDY*EAA*COO etkileşimi en uygun modele girmiş ve istatistiksel olarak anlamlı bulunmuştur.

En uygun modele ilişkin etkileşimlerin, en uygun model bulunmadan önceki parametre tahminleri Tablo-6'da verilmiştir. Tabloda görüldüğü gibi yalnızca AED*BED etkileşiminde (AED_OY+D)-(BED_OY+D), (AED_İlk)-(BED_OY+D) ve (AED_İlk)-(BED_İlk) %5 anlam düzeyinde önemli, diğer parametrelerin ise önemsiz olduğu bulunmuştur.

Tablo-6: En uygun modelin, en uygun model bulunmadan önceki kısmi ilişkilerinin (etkileşimler) standart hata, z ve p değerleri.

Kısmi İlişkiler (Etkileşimler)	Parametre	Parametrelerin Açıklaması	Standart Hata	z	P
ADY*EAA*COO	1	(ADY_Köy)-(EAA_Var)-(COO_Var)	0,0608	-0,73559	>0,05
	2	(ADY_İlçe)-(EAA_Var)-(COO_Var)	0,0609	+0,40329	>0,05
ADY*BDY*BED	1	(ADY_Köy)-(BDY_Köy)-(BED_OY+D)	0,1768	-1,89149	>0,05
	2	(ADY_Köy)-(BDY_Köy)-(BED_İlk)	0,1576	+1,87136	>0,05
	3	(ADY_Köy)-(BDY_Köy)-(BED_Orta)	0,1651	+0,14104	>0,05
	4	(ADY_Köy)-(BDY_Köy)-(BED_Lise)	0,1654	+0,44228	>0,05
	5	(ADY_Köy)-(BDY_İlçe)-(BED_OY+D)	0,1806	+0,94510	>0,05
	6	(ADY_Köy)-(BDY_İlçe)-(BED_İlk)	0,1701	-0,83399	>0,05
	7	(ADY_Köy)-(BDY_İlçe)-(BED_Orta)	0,1742	-0,66852	>0,05
	8	(ADY_Köy)-(BDY_İlçe)-(BED_Lise)	0,1761	+0,24756	>0,05
	9	(ADY_İlçe)-(BDY_Köy)-(BED_OY+D)	0,1792	+1,23370	>0,05
	10	(ADY_İlçe)-(BDY_Köy)-(BED_İlk)	0,1692	-1,56783	>0,05
	11	(ADY_İlçe)-(BDY_Köy)-(BED_Orta)	0,1714	+0,34700	>0,05
	12	(ADY_İlçe)-(BDY_Köy)-(BED_Lise)	0,1714	-0,48094	>0,05
	13	(ADY_İlçe)-(BDY_İlçe)-(BED_OY+D)	0,1795	-1,24092	>0,05
	14	(ADY_İlçe)-(BDY_İlçe)-(BED_İlk)	0,1645	+1,37295	>0,05
	15	(ADY_İlçe)-(BDY_İlçe)-(BED_Orta)	0,1672	+0,27735	>0,05
	16	(ADY_İlçe)-(BDY_İlçe)-(BED_Lise)	0,1690	+0,12379	>0,05
BDY*EAA*COO	1	(BDY_Köy)-(EAA_Var)-(COO_Var)	0,0602	+0,67440	>0,05
	2	(BDY_İlçe)-(EAA_Var)-(COO_Var)	0,0611	-0,49241	>0,05
AED*BED	1	(AED_OY+D)-(BED_OY+D)	0,1707	+3,06933	<0,01
	2	(AED_OY+D)-(BED_İlk)	0,1702	+0,31360	>0,05
	3	(AED_OY+D)-(BED_Orta)	0,1776	-0,50071	>0,05
	4	(AED_OY+D)-(BED_Lise)	0,1799	-0,96260	>0,05
	5	(AED_İlk)-(BED_OY+D)	0,1806	-2,23957	<0,05
	6	(AED_İlk)-(BED_İlk)	0,1487	+4,24410	<0,001
	7	(AED_İlk)-(BED_Orta)	0,1595	+0,48718	>0,05
	8	(AED_İlk)-(BED_Lise)	0,1609	+0,22033	>0,05
	9	(AED_Orta)-(BED_OY+D)	0,1839	-0,11927	>0,05
	10	(AED_Orta)-(BED_İlk)	0,1701	-0,36064	>0,05
	11	(AED_Orta)-(BED_Orta)	0,1663	+1,40694	>0,05
	12	(AED_Orta)-(BED_Lise)	0,1727	-0,01008	>0,05
	13	(AED_Lise)-(BED_OY+D)	0,1831	-1,04268	>0,05
	14	(AED_Lise)-(BED_İlk)	0,1704	-1,74999	>0,05
	15	(AED_Lise)-(BED_Orta)	0,1698	-0,42535	>0,05
	16	(AED_Lise)-(BED_Lise)	0,1652	+1,28550	>0,05
ADY*AED	1	(ADY_Köy)-(AED_OY+D)	0,1219	+1,44340	>0,05
	2	(ADY_Köy)-(AED_İlk)	0,1148	+0,75909	>0,05
	3	(ADY_Köy)-(AED_Orta)	0,1229	-0,52521	>0,05
	4	(ADY_Köy)-(AED_Lise)	0,1205	-0,90859	>0,05
	5	(ADY_İlçe)-(AED_OY+D)	0,1245	-0,18886	>0,05
	6	(ADY_İlçe)-(AED_İlk)	0,1156	-0,04635	>0,05
	7	(ADY_İlçe)-(AED_Orta)	0,1215	+0,51476	>0,05
	8	(ADY_İlçe)-(AED_Lise)	0,1204	-0,37890	>0,05
AED*EAA	1	(AED_OY+D)-(EAA_Var)	0,0879	+1,25268	>0,05
	2	(AED_İlk)-(EAA_Var)	0,0817	-0,65991	>0,05
	3	(AED_Orta)-(EAA_Var)	0,0868	+0,35705	>0,05
	4	(AED_Lise)-(EAA_Var)	0,0848	-1,20463	>0,05

2. EN UYGUN LOG-LİNEER MODELİN ETKİLEŞİMLERİNİN HOMOJENİTE ANALİZİ

2.1. ADY*EAA*COO Etkileşimi İçin HA

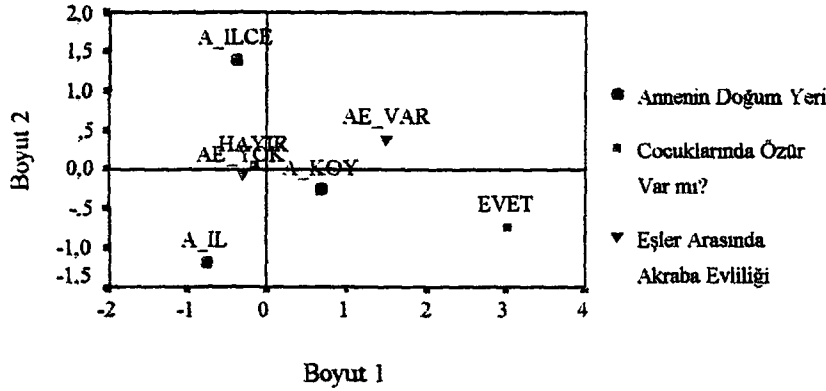
Geriye doğru aşamalı LLA sonucunda bulunan modelin ADY*EAA*COO etkileşimine ilişkin çoklu çapraz tablo, Tablo-7’de verilmiştir. Sözkonusu etkileşime ilişkin çapraz tablonun HA sonucunda elde edilen grafiksel gösterimi ise Şekil-2’deki gibidir.

Tablo-7:ADY*EAA*COO etkileşimine ilişkin çoklu çapraz tablo.

Annenin Doğum Yeri	Eşler Arasında Akraba Evliliği				Toplam
	Var		Yok		
	Çocuklarında Özür Olup Olmadığı		Çocuklarında Özür Olup Olmadığı		
	Var	Yok	Var	Yok	
Köy	17	87	18	319	441
İlçe	4	43	2	243	292
İl	1	22	4	218	245
Toplam	22	152	24	780	978

Not: 22 aile ilgili değişkenlere cevap vermemiştir.

Kategori Koordinatları



Şekil-2: ADY*EAA*COO etkileşimi için HA.

Sekil-2 incelendiğinde anne köy doğumlu ise eşlerin akraba evliliği yapma

akraba evliliği yapmama eğiliminde oldukları ve çocuklarında özür görülmediği gözlenmektedir. Eşlerin akraba evliliği yapmasını etkileyen neden annenin doğum yeri olup, çocuklarında özür görülmesini etkileyen neden ise eşlerin akraba evliliği yapmalarıdır. Şekil-2'nin birinci eksenini (boyutu) olayın %56.28'ini; ikinci eksenini ise %43.72'sini açıklamaktadır.

2.2. BDY*EAA*COO Etkileşimi İçin HA

Geriye doğru aşamalı LLA sonucunda bulunan modelin BDY*EAA*COO etkileşimine ilişkin çoklu çapraz tablo, Tablo-8'de verilmiştir. Söz konusu etkileşime ilişkin çapraz tablonun HA sonucunda elde edilen grafiksel gösterim Şekil-3'deki gibidir.

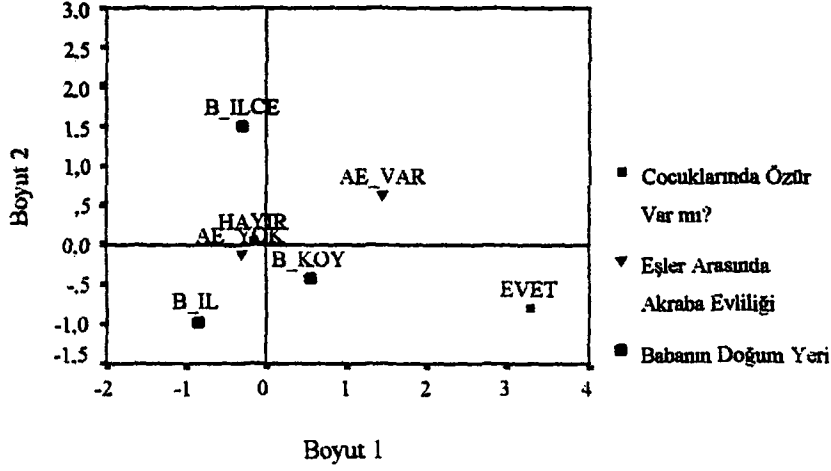
Tablo-8: BDY*EAA*COO etkileşimine ilişkin çoklu çapraz tablo.

Babanın Doğum Yeri	Eşler Arasında Akraba Evliliği				Toplam
	Var		Yok		
	Çocuklarında Özür Olup Olmadığı		Çocuklarında Özür Olup Olmadığı		
	Var	Yok	Var	Yok	
Köy	19	78	18	366	481
İlçe	2	49	3	220	274
İl	0	22	3	181	206
Toplam	21	149	24	767	961

Not: 39 aile ilgili değişkenlere cevap vermemiştir.

Şekil-3 incelendiğinde, baba köy doğumlu ise eşlerin akraba evliliği yapma eğiliminde olduğu ve çocuklarında özür görüldüğü, bunun yanı sıra baba ilçe ya da il doğumlu ise eşlerin akraba evliliği yapmama eğiliminde olduğu ve çocuklarında özür olmadığı görülmektedir. Birinci boyut olayın %55.52'sini; ikinci boyut ise %44.48'ini açıklamaktadır. ADY*EAA*COO etkileşimi ile BDY*EAA*COO etkileşimleri grafiksel olarak benzer sonuçlar vermiştir. Annenin doğum yeri ile babanın doğum yeri akraba evliliğini etkilemekte ve dolayısıyla çocuklarında özür görülmesine sebebiyet vermektedir.

Kategori Koordinatları



Şekil-3: BDY*EAA*COO etkileşimi için HA.

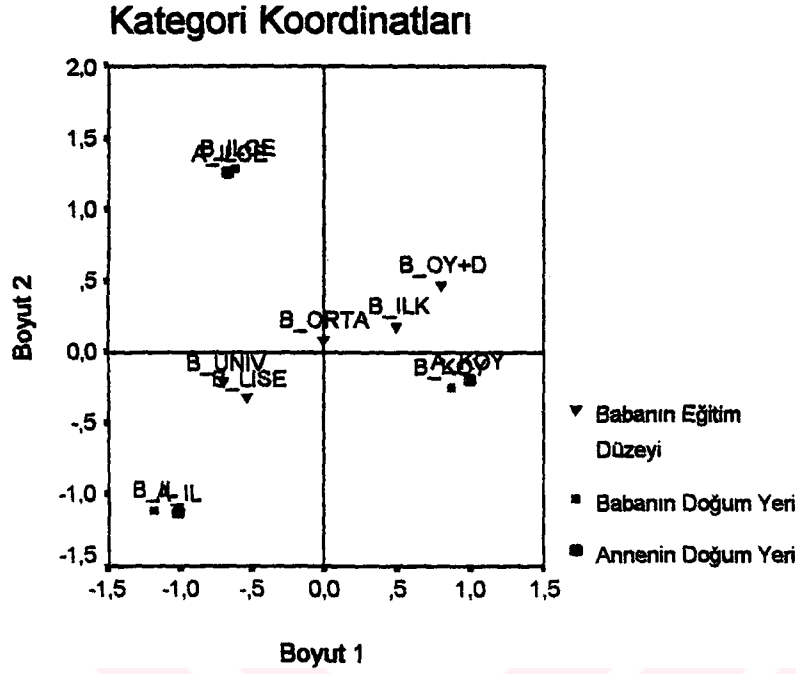
2.3. ADY*BDY*BED Etkileşimi İçin HA

Geriye doğru aşamalı LLA sonucunda bulunan modelin ADY*BDY*BED etkileşimine ilişkin çoklu çapraz tablo, Tablo-9'da verilmiştir. Söz konusu etkileşime ilişkin çapraz tablonun HA sonucunda elde edilen grafiksel gösterim Şekil-4'teki gibidir.

Tablo-9: ADY*BDY*BED etkileşimine ilişkin çoklu çapraz tablo.

Babanın Eğitim Düzeyi	Babanın Doğum Yeri									Toplam
	Köy			İlçe			İl			
	Annenin Doğum Yeri			Annenin Doğum Yeri			Annenin Doğum Yeri			
	Köy	İlçe	İl	Köy	İlçe	İl	Köy	İlçe	İl	
Okuryazar+ okuryazar değil	16	3	1	2	6	0	1	0	1	30
İlkokul	227	10	22	16	95	7	4	10	40	431
Orta okul	50	5	3	3	28	8	6	3	18	124
Lise	53	6	7	2	33	5	1	19	41	167
Yüksekokul	39	7	21	10	39	18	6	10	46	196
Toplam	385	31	54	33	201	38	18	42	146	948

Not: 52 aile ilgili değişkenlere cevap vermemiştir.



Şekil-4: ADY*BDY*BED etkileşimi için HA.

Şekil-4’de görüldüğü gibi, anne ve babanın doğum yerleri aynıdır. Anne ve babanın doğum yeri köy ise babanın eğitim düzeyi okuryazar değil, okuryazar veya ilkokul mezunudur. Anne ve babanın doğum yeri ilçe ya da il ise, baba orta okul, lise ya da üniversite mezunudur. Birinci boyut olayın %53.94’ünü; ikinci boyut ise %46.06’sını açıklamaktadır.

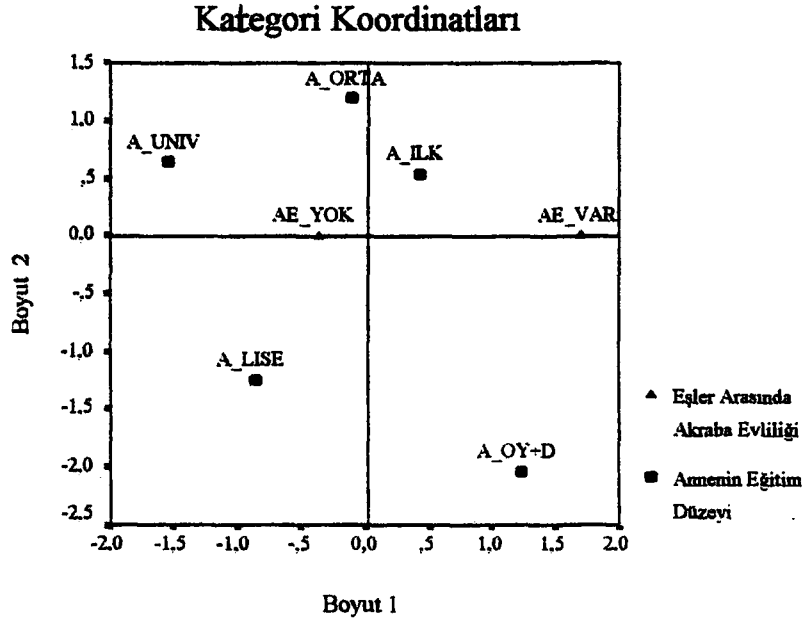
2.4. AED*EAA Etkileşimi İçin HA

Geriye doğru aşamalı LLA sonucunda bulunan modelin AED*EAA etkileşimine ilişkin çapraz tablo, Tablo-10’da verilmiştir. Sözkonusu etkileşime ilişkin çapraz tablonun HA sonucunda elde edilen grafiksel gösterim Şekil-5’deki gibidir.

Tablo-10: AED*EAA etkileşimine ilişkin çapraz tablo.

Eşler Arasında Akraba Evliliği	Annenin Eğitim Düzeyi					Toplam
	Okur-yazar+ okur-yazar değil	İlkokul	Orta okul	Lise	Yüksekokul	
Var	28	115	13	17	1	174
Yok	64	406	66	183	93	812
Toplam	92	521	79	200	94	986

Not: 14 aile ilgili değişkenlere cevap vermemiştir.



Şekil-5: AED*EAA etkileşimi için HA.

Şekil-5'e göre anne okuryazar değil, okuryazar ya da ilkokul mezunu ise akraba evliliği yapma eğiliminde olduğu; bunun yanı sıra anne orta okul, lise ya da üniversite mezunu ise akraba evliliği yapmama eğiliminde olduğu gözlenmiştir. Dolayısıyla akraba evliliği yapmayı etkileyen diğer bir neden de annenin eğitim düzeyidir. Birinci boyut olayın %54.79'unu; ikinci boyut ise %45.21'ini açıklamaktadır.

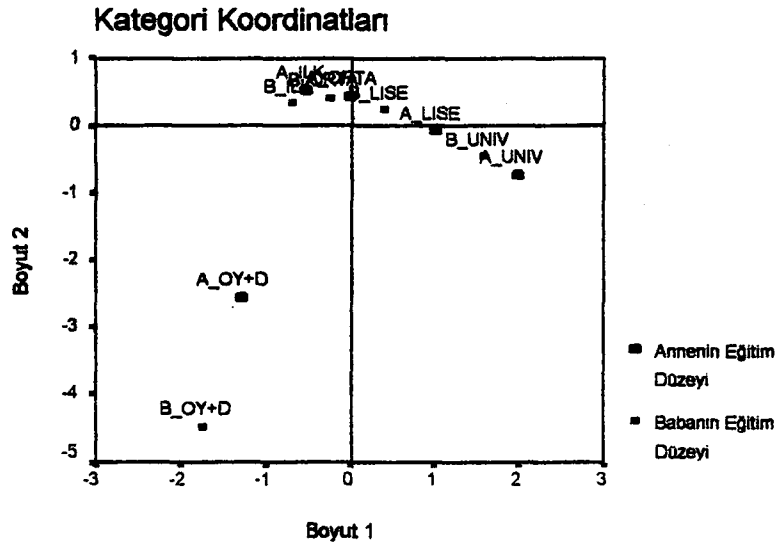
2.5. AED*BED Etkileşimi İçin HA

Geriye doğru aşamalı LLA sonucunda bulunan modelin AED*BED etkileşimine ilişkin çapraz tablo, Tablo-11'de verilmiştir. Söz konusu etkileşime ilişkin çapraz tablonun HA sonucunda elde edilen grafiksel gösterim Şekil-6'daki gibidir.

Tablo-11: AED*BED etkileşimine ilişkin çapraz tablo.

Babanın Eğitim Düzeyi	Annemin Eğitim Düzeyi					Toplam
	Okur-yazar+ okur-yazar değil	İlkokul	Orta okul	Lise	Yüksekokul	
Okur-yazar+ okur-yazar değil	29	1	0	0	0	30
İlkokul	45	348	24	17	1	435
Orta okul	7	67	29	20	1	124
Lise	4	68	14	76	6	168
Yüksekokul	1	25	11	84	85	206
Toplam	86	509	78	197	93	963

Not: 37 aile ilgili değişkenlere cevap vermemiştir.



Şekil-6: AED*BED etkileşimi için HA.

Şekil-6 incelendiğinde, birinci boyut açısından bakıldığında anne ve babanın eğitim düzeylerinin ilkokuldan başlayarak üniversiteye doğru sıralandığı, fakat okur-yazar olmayan ve okur-yazar olanların ise bu sıradan farklı yerde olduğu görülmekte; ikinci boyut açısından bakıldığında ise anne ve babanın ilkokul, orta okul, lise ve üniversite mezunu olma açısından birbirleriyle benzerlik gösterdiği; okur-yazar olma ve okur-yazar olmama açısından ise ilkokul, orta okul, lise ve yüksek okul mezunu olma durumuna göre benzerlik göstermedikleri görülmektedir. Birinci boyut toplam varyansın %52.89'ünü; ikinci boyut ise %47.11'ini açıklamaktadır.

2.6. ADY*AED Etkileşimi İçin HA

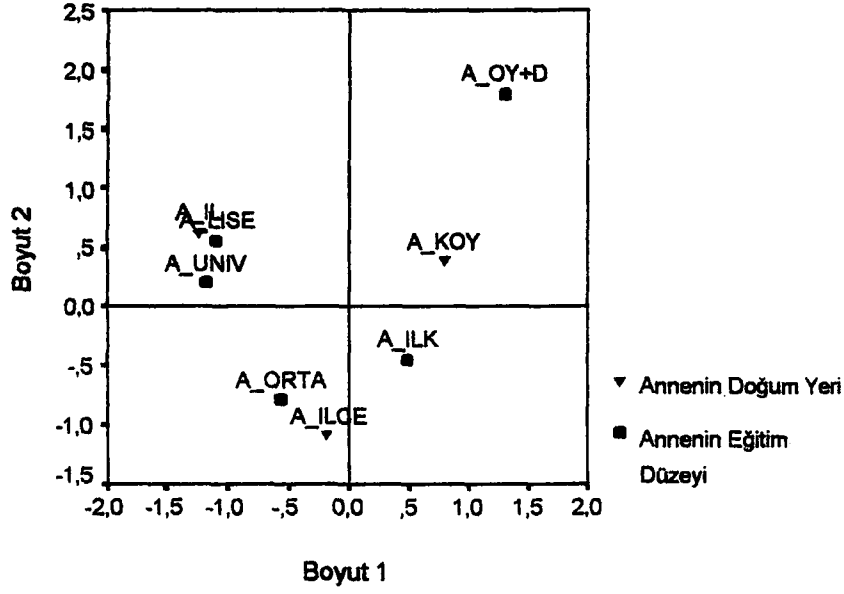
Geriye doğru aşamalı LLA sonucunda bulunan modelin ADY*AED etkileşimine ilişkin çapraz tablo, Tablo-12'de verilmiştir. Sözkonusu etkileşime ilişkin çapraz tablonun HA sonucunda elde edilen grafiksel gösterim Şekil-7'deki gibidir.

Tablo-12: ADY*AED etkileşimine ilişkin çapraz tablo.

Annenin Doğum Yeri	Annenin Eğitim Düzeyi					Toplam
	Okur-yazar+ okur-yazar değil	İlkokul	Orta okul	Lise	Yüksek okul	
Köy	67	281	27	49	21	445
İlçe	21	153	26	64	31	295
İl	5	86	27	86	42	246
Toplam	93	520	80	199	94	986

Not: 14 aile ilgili değişkenlere cevap vermemiştir.

Kategori Koordinatları



Şekil-7: ADY*AED etkileşimi için HA.

Şekil 7'ye göre; anne köy doğumlu ise okuryazar değil, okuryazar ya da ilkokul mezunu olduğu; anne ilçe doğumlu ise orta okul; il doğumlu ise lise ya da üniversite mezunu olduğu görülmektedir. Birinci boyut olayın %56.94'ünü; ikinci boyut ise %43.06'sını açıklamaktadır.

TARTIŞMA VE SONUÇ

Günümüzde tıp ve biyoloji alanında yapılan araştırmalarda incelemeye alınan birimlere ilişkin çok sayıda kategorik değişkenlerin gözlenmesi söz konusudur. Bu tür araştırmalarda amaç, gözlem sonucu elde edilen kategorik değişkenlerin çeşitli çapraz tablolar halinde düzenlenmesi ve bu tablolar üzerinde uygulanacak istatistiksel yöntemlerle güvenilir sonuçlara ulaşmaktır.

Birbirleri ile ilişkili olduğu bilinen kategorik yapıdaki çok sayıda değişkenin hangilerinin birbirleri ile ilişkili olduğunu belirlemek ve bu değişkenleri analiz etmek güvenilir sonuçlara ulaşmayı sağlar. Bu amaçla uygulanan en iyi analiz yöntemi LLA yöntemidir. LLA değişkenler arasındaki bağımlılığın araştırılmasında yorumlanması kolay olan modellerden oluşur.

Çapraz tabloların bağımsızlık analizlerinde satır ve sütunları birbirleriyle karşılaştırmak için satır ve sütun yüzdeleri kullanılabilir, fakat sınıf sayıları artış gösterdikçe sınıfların ait olduğu satır ve sütunların birbiriyle karşılaştırılmaları güçleşir hatta bazen imkansızlaşır. Değişkenlere ait satır ve sütun sayısı fazla olan kontenjans tablolarının analizinde ki-kare bağımsızlık analizleri yetersiz kalmaktadır. Satır ve sütun sayısına kısıt getirmeyen, değişkenlerin alt kategorilerinin kolayca yorumlanabilmesini sağlayan CA ya da HA tercih edilmektedir.

HA, CA'nın genişletilmiş şekli olup değişkenlerin iki ya da daha çok setleri arasındaki ilişkileri analiz etmek için kullanılan çok değişkenli bir istatistiksel yöntemdir. Bu araştırma, LLA ile HA'nın birbirlerinin tamamlayıcısı olarak kullanılıp kullanılmayacağını belirlemek amacıyla yapılmıştır.

Konuyla ilişkili olan literatür taramalarında; Tenenhaus ve Young (1985), yapmış oldukları bir çalışmada kategorik çok değişkenli verileri ölçmek için çeşitli metotları tartışmışlar ve bunların arasında HA'nın çoğu kez önerilen bir teknik olduğunu ve kategorik yapıdaki değişken matrisiyle bir nesnenin analizinin yapılmasının uygun bir yol olduğunu ifade etmişlerdir (17). de Leeuw (1985), CA tekniğinin iki yönlü tablolarda bağımsız olan modellerden hata terimlerinin analizinde faydalı olabildiğini ifade etmişlerdir (45). Goodman (1986), CA modellerinin nasıl genel uygunluk analiz yaklaşım yönteminin bir genişletilmiş hali olarak sunulabileceğini ve log-bilineer modellerin nasıl

genel log-lineer model yaklaşımının bir genişletilmiş hali olacağını açıklamıştır (13). van der Heijden ve Worsley (1988) ve van der Heijden ve ark. (1989) CA ile LLA yöntemini karşılaştırmışlar ve bunun sonucu olarak CA'nın log-lineer modellerde artıkların analizi için kullanılabileceğini belirtmişlerdir (15, 16). Türe (1993), CA'nın çok sayıda satır ve sütundan oluşan çapraz tabloların analizinde uygun bir yöntem olduğunu belirtmiştir (8).

Bu çalışmada, Denizli ilinde 1000 aileye uygulanan akraba evliliği ile ilgili bir araştırmanın akraba evliliği yapmayı ve özürlü çocuğa sahip olmayı etkileyebilecek olan kategorik yapıdaki "annenin eğitim durumu", "annenin doğum yeri", "babanın eğitim durumu", "babanın doğum yeri", "eşler arasında akraba evliliği", "çocuklarda özür olup olmadığı" değişkenleri ele alınmış ve bu değişkenlere önce geriye doğru aşamalı LLA, elde edilen modelin bileşenlerine de HA uygulanmıştır. Geriye doğru aşamalı LLA sonucunda en fazla üçlü etkileşimler istatistiksel olarak anlamlı bulunmuştur. İlgili değişkenlerin alt kategorileri arasındaki ilişkinin net bir şekilde görülebilmesi için en uygun modelin bileşenlerine HA uygulanmış ve bileşenlerin alt kategorilerinin birbirleri ile olan etkileşimleri grafiksel olarak yorumlanmıştır. Bunun sonucunda, anne ve babanın doğum yerleri ve eğitim düzeylerinin eşlerin akraba evliliği yapmasını etkilediği ve akraba evliliğinin ise çocuklarda özür görülmesine sebebiyet verdiği görülmüştür.

Başaran (1995), ülkemizde doğudan batıya doğru gidildikçe akraba evliliği oranında bir azalma olduğuna dikkat çekmektedir. Türkiye'deki akraba evliliğinin toplumun sosyo-ekonomik yapısı, dini inançları, töreleri, yerleşme yerinin büyüklüğü ve coğrafik konum gibi faktörlerin etkisinde olduğunu ve zamana bağlı olarak farklılık gösterdiğini belirtmektedir (46).

Ulusoy ve Tunçbilek (1987), Türkiye'de akraba evliliklerinin yaygın bir evlilik biçimi olmakla birlikte, ülkenin az gelişmiş ve gelişmiş bölgeleri (kır ve kent yerleşim yerleri) arasında ve değişik eğitim düzeylerinde farklı yoğunlukta olduğunu; gelişmiş bölgelerde (kentlerde) yüksek eğitim düzeyinde akraba evliliklerinin azaldığını, kadın ve erkeklerin genel olarak doğum yerleri ve eğitim düzeylerine bakıldığında; doğum yerleri daha az gelişmiş bölge (köy) ve eğitim düzeyleri ilk ve daha az ise akraba evliliği yapma eğiliminin artış gösterdiğini, eşlerin birlikte eğitim durumlarına bakıldığında, akraba evliliği yapmanın eğitimle ters ilişkili olduğunu belirtmişlerdir. Eğitimli çiftlerin daha az

akraba evliliği yapma eğilimine karşı, eğitimsiz çiftlerin daha çok akraba evliliği yapma eğiliminde olduğunu, doğum yerleri açısından bakıldığında yine benzer sonuçlarla karşılaşıldığını belirtmişlerdir. İlçe ve il doğumlu olanların daha az akraba evliliği yapma eğilimine karşı, daha az gelişmiş bölge doğumluların ise daha çok akraba evliliği yapma eğilimi gösterdiğini belirtmişlerdir (47). Bu sonuçlar ile çalışmamızda elde ettiğimiz sonuçlar paralellik göstermektedir.

Araştırmamız sonucu elde ettiğimiz sonuçlar yöntem olarak, Tenenhaus ve Young (1985); van der Heijden ve Worsley (1988); van der Heijden ve ark. (1989)'nın yapmış oldukları çalışmaların sonuçları ile paralellik göstermektedir (15,16,17).

LLA'nın çok sayıda değişkenin hangilerinin birbirleri ile ilişkili olduğunu belirlemede iyi bir yöntem olduğu, fakat bunun yanında birbirleri ile ilişkili olan değişkenlerin alt kategorilerinin hangilerinin ilişkili olduğunu belirlemede çok yetersiz hatta tutarsız olduğu gözlenmiştir.

HA çok sayıda kategorik değişkenin hangilerinin birbirleri ile ilişkili olduğunu belirleyememesi fakat bunun yanısıra birbirleri ile ilişkili olduğu bilinen iki ya da ikiden fazla değişkenin alt kategorilerinin hangilerinin birbirleri ile ilişkili olduğunu grafiksel olarak ortaya koyan bir yöntem olması nedeniyle LLA ile HA'nın birbirlerini bütünleyen analiz yöntemleri olduğu gözlenmiştir.

Kısaca bu iki yöntemin birlikte kullanılabilceği ve sonuçların daha güvenilir ve yorumlamayı kolaylaştırıcı olacağı görülmüştür. Sonuç olarak istatistiksel analiz yöntemlerini uygulayan araştırmacıların kategorik verilerin analizinde LLA'nın rakam yığınları ile uğraşmaları yerine, LLA'nın en uygun modeline HA uygulamaları ve sonuçları grafiksel gösterim ile yorumlamalarının daha kolay, net ve güvenilir olacağı görüşüne varılmıştır.

ÖZET

Biyolojik ve tıbbi denemelerin çoğunda incelenen değişkenlerin nitel olması nedeniyle, elde edilen kesikli değişkenlerin analizinde parametrik olmayan yöntemler kullanılmaktadır. Parametrik olmayan çok değişkenli istatistiksel analiz yöntemlerinden log-lineer analiz ve homojenite analizi değişken sayısının ikiden fazla olduğu çok değişkenli kontenjans tablolarında değişkenler arasındaki karmaşık yapıda olan ilişkilerin daha kolay anlaşılabilmesine olanak tanıyan analiz yöntemleridir.

Bu çalışmada log-lineer analiz ile homojenite analizinin birbirinin tamamlayıcısı olup olmadıkları belirlenmeye çalışılmıştır. Akraba evliliğini etkileyen faktörlerin belirlenmesi için 1996 yılında Denizli il merkezinde yaşayan 1000 aileye ilişkin bir araştırmada ailelere sorulan sorulardan akraba evliliğini etkileyebilecek değişkenlere log-lineer analiz yöntemi uygulanmıştır. Geriye doğru aşamalı log-lineer analiz yöntemi ile $\ln(f_{ijklmn}) = \mu + \alpha_{ACE} + \alpha_{ADF} + \alpha_{DCE} + \alpha_{BF} + \alpha_{AB} + \alpha_{BC}$ modeli en uygun model olarak bulunmuştur. Bu en uygun modele homojenite analizi uygulanmıştır. Bunların sonucunda elde edilen analiz sonuçları incelendiğinde; düşük eğitim düzeyli, köy doğumlu olan eşlerde akraba evliliğine karşı eğilim olduğu bulunmuştur. Bunun tersi durum ise il ya da ilçe doğumlu olan eşler için elde edilmiştir. Akraba evliliğinin özürlü çocuk sahibi olmada etken olduğu görülmüştür. Anne ve baba köy doğumluysa ve bunun yanısıra eğitim düzeyleri okuryazar değil, okuryazar ya da ilkokul ise bu durumda akraba evliliği yapma eğiliminde, bunun aksine ilçe ya da il doğumlu iseler ve eğitim düzeyleri orta okul, lise ya da üniversite ise bu durumda akraba evliliği yapmama eğiliminde oldukları görülmüştür.

Bu çalışma sonucunda log-lineer ve homojenite analizinin birlikte kullanılabileceği ve sonuçların daha güvenilir ve yorumlamayı kolaylaştırıcı olacağı görülmüştür. İstatistiksel analiz yöntemlerini uygulayan araştırmacıların kategorik verilerin analizinde log-lineer analizinin rakam yığınları ile uğraşmaları yerine, log-lineer analizinin en uygun modeline homojenite analizini uygulamaları ve sonuçları grafiksel gösterim ile yorumlamalarının daha kolay, objektif ve güvenilir olacağı görüşüne varılmıştır.

COMPARISON OF HOMOGENEITY AND LOG-LINEAR ANALYSES FROM CONTINGENCY TABLES

SUMMARY

In the most of biologic and medical studies, because of the fact that variables studied are quantitative, non-parametric methods have been used in the analysis of the obtained discontinuous variables. Of the non-parametric multivariate statistical methods log-linear and homogeneity analyses provide opportunity for understanding the complex relationship among the variables in the contingency tables with more than two variables.

In this study, we investigated the possibility that whether log-linear and homogeneity analyses are complementary to each other. The study was carried out on 1000 families in 1996. Log-linear analysis was used to determine the factors which affect the consanguineous marriage among the residents of Denizli City Center. $\ln(f_{ijklmn}) = \mu + \alpha_{ACE} + \alpha_{ADF} + \alpha_{DCE} + \alpha_{BF} + \alpha_{AB} + \alpha_{BC}$ was found to be the most appropriate model by phasic-retrograde log-linear analysis method. On application of homogeneity to analyse the concerned appropriate model and investigation of the obtained results, a low education level and high tendency consanguineous marriage were found among rural-born pairs. A reverse was obtained for district or province-born pairs. Consanguineous marriage was found to cause congenital disorders. When the parents were rural-born their educational level was either illiterate or literate and/or primary school graduate with a tendency to consanguineous marriage. No consanguineous marriage was observed when the parents were district or province-born and graduates of either secondary, high school or university.

In conclusion, these results revealed applicability of combined log-linear and homogeneity analyses. This renders the results more reliable and facilitates their interpretation. Instead of being dealing with massive numerical data, log-linear analysis enables determining the most appropriate model, application of homogeneity analysis and presentation of the data grafically in a simple, objective and reliable way.

KAYNAKLAR DİZİNİ

1. Kotze, V.W.: The log-linear model and application to multiway contingency tables, Cambridge Univ., 1982.
2. Ku, H. H, Varner, R. N., Kullback, S.: On the analysis of multidimensional contingency tables, JASA, **66**, 55-63, 1971.
3. SPSS Release 6.1: Categories handbook, 1994.
4. SPSS Release 6.1: Advanced Statistics, 1994.
5. Haberman, J.S., Gilula, Z.: The analysis of multivariate contingency tables by restricted canonical and restricted association models, JASA, **83**, 770-771, 1988.
6. Benedetti, J. K., Brown, N. B.: Strategies for the selection of log-linear models, Biometrics, **34**, 680-686, 1978.
7. Rao, J. N. K., Scott, A. J.: On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data, Annals of Statistics, **12**, 157-170, 1984.
8. Türe, M.: Frekans tablolarında satır, sütun ve birim sayısının uyum analizi sonuçları üzerine etkisinin araştırılması, (Yayınlanmamış doktora tezi), 1993.
9. Agresti, A.: Categorical data analysis, John Wiley and Sons, New York, 1990.
10. Christensen, R.: Log-linear models, library of congress cataloging in publication data, 1990.
11. Cormack, R. M.: Log-linear models for capture-recapture, Biometrics, **45**, 221-230, 1989.
12. Dobson, A. J.: An introduction generalized linear models, Chapman and Hall, New York, 1990.
13. Goodman, Leo A.: Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables. Internat. Statist. Rev., **54**, 243-270, 1986.
14. de Leeuw and van der Heijden, P. G. M.: Correspondence analysis of incomplete contingency tables. Psychometrika, **53(2)**, 223-233, 1988.
15. van der Heijden, P.G.M. and Worsley, K. J.: Comment on "correspondence analysis used complementary to log-linear analysis". Psychometrika, **53(2)**, 287-291, 1988.
16. van der Heijden, P. G. M. de Falguerolles, A. and de Leeuw, J.: A combined approach to contingency table analysis using correspondence analysis and log-linear analysis, Applied Statistics, **38(2)**, 249-292, 1989.
17. Tenenhaus, M., Young, F.W.: An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data, Psychometrika, **50**, 91-119, 1985.

18. van der Burg, E., de Leeuw J.: Homogeneity analysis with k sets of variables: An alternating least squares method with optimal scaling features, *Psychometrica*, **53(2)**, 177-197, 1988.
19. Daniel, W.W.: Applied non-parametric statistics. Houghton Mifflin Company, Dallas, 1978.
20. Aickin, M.: Linear statistical analysis of discrete data. John Wiley and Sons, New York, 1983.
21. Agresti, A.: Analysis of ordinal categorical data. John Wiley and Sons, New York, 1984.
22. Simonoff, J. S., Chih-Ling Tsai: Higher order effects in log-linear models and log non linear models for contingency tables with ordered categories, *App. Statist.*, **40(3)**, 449-458, 1991.
23. Özdamar, K.: Paket programlar ile istatistiksel veri analizi I. Anadolu Üniversitesi Yayınları, No:1001, Eskişehir, 1997.
24. Andersen, E.B.: The statistical analysis of categorical data, Springer-Verlag, Berlin, 1990.
25. Haberman, S. J.: Log-linear fit for contingency tables (Algoritm AS 57) *App. Statist.*, **21**, 218-225, 1972.
26. Everitt, S. B., Dunn, G.: Applied multivariate data analysis, John Wiley and Sons, New York, 1991.
27. Yılmaz, V.: Türkiye'deki intiharlara ilişkin kategorik verinin log-linear modeller ile analizi, (Yayınlanmamış doktora tezi), Eskişehir, 1996.
28. Tatlıdil, H.: Uygulamalı çok değişkenli istatistiksel analiz, Ankara, 1992.
29. Hagenauers, J. A.: Log-linear models with latent variables, Sage Publications, Newbury Park, 1993.
30. Knoke, D., Burke, P.: Log-linear models, Sage, Beverly Hills 07-20, 1980.
31. Moran, A., Engelman, L., Stephen, E., and Gerald, G.F.: BMDP Inc. California, 1988.
32. Gange, J.S.: Generating multivariate categorical variates using the iterative proportional fitting algorithm, *The American Statistician*, **49(2)**, 134-138, 1995.
33. Gökhale, D. V.: Analysis of log-linear models, *App. Stat.*, **3**, 371-376, 1972.
34. van der Heijden, P.G.M.: Correspondence analysis of transition matrices. *KM*, **19**, 49-72, 1985.
35. van der Heijden, P.G.M. and de Leeuw, J.: Correspondence analysis, with special attention to the analysis of panel data and event history data. *Sociological Methodology*, 43-87, 1989.
36. Pack, P. and Jolliffe, I. T.: Influence in correspondence analysis, *Appl. Statist.*, **41(2)**, 365-380, 1992.

37. Hawkins, D.M.: Topics in applied multivariate analysis. Cambridge University Press, USA, 1982.
38. Lebart, L., Monriveau, A. and Warwick, K. M.: Multivariate descriptive statistical analysis. John Wiley and Sons, New York, 1984.
39. Jambu, M., Lebeaux, M.O.: Cluster analysis and data analysis. Nord-Holland Pub., Amsterdam, 1986.
40. van Rijckevorsel, J., de Leeuw J.: Component and correspondence analysis, John Wiley and Sons, New York, 1988.
41. Barnett, V.: Interpreting multivariate data, John Wiley and Sons, New York, March, 1980.
42. van der Heijden, P. G. M.: Transition matrices, model fitting and correspondence analysis. Data Analysis and Informatics IV, Amsterdam, 1986.
43. van der Heijden, P. G. M. and Escofier, B.: Multiple corresponding analysis with missing data. Institute National de Recherche en Informatique et en Automatique, 1988.
44. van der Heijden, P. G. M., de Vries, H. and van Hoof, J.A.R.A.M.: Correspondence analysis of transition matrices, with special attention to missing entires and asymmetry. Anim. Behav., 40, 49-64, 1990.
45. de Leeuw, J.: Correspondence analysis used complementary to log-linear analysis. Psychometrika, 50(4), 429-447, 1985.
46. Başaran, N.:Genetik hastalıkları doğumöncesi tam bülteni, 4-5, 1995.
47. Ulusoy, M., Tunçbilek, E.:Türkiye’de akraba evlilikleri ve çocuk ölümlerine etkisi, Nüfusbilim derg./Türk. J. Population Study, 9, 7-26, 1987.

TEŐEKKÖR

Bu tezin hazırlanmasında deęerli grüş ve dűőünceleriyle bana her aőamada yön veren danıőman hocam sayın Yrd. Doę. Dr. Mevlüt TÜRE'ye teőekkürlerimi sunarım.

Ayrıca alıőmalarım sırasında her türlü desteklerini esirgemeyen aileme de teőekkürlerimi sunmayı bir bor bilirim.



ÖZGEÇMİŞ

1972 yılında Adana'da doğdu. İlkokulu 1983 yılında ve ortaöğrenimini 1989 yılında tamamladı. 1991 yılında Eskişehir Fen Edebiyat Fakültesi İstatistik Bölümü'ne girdi. Bu bölümden 1995 yılında mezun oldu ve aynı yıl Pamukkale Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı'nda araştırma görevlisi olarak göreve başladı. 1996 yılının Eylül ayında ise Pamukkale Üniversitesi Biyoistatistik Anabilim Dalı'nda yüksekisansa başladı. Halen Pamukkale Üniversitesi Biyoistatistik Anabilim Dalı'nda araştırma görevlisi olarak görevini sürdürmektedir.



T.C. YÜKSEKÖĞRETİM KURULU
DOKÜMANTASYON MERKEZİ