

T. C.
YÜZÜNCÜ YIL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

**DOĞRUSAL OLMAYAN KANONİK KORELASYON ANALİZİ
VE SAĞLIK ALANINDA UYGULAMASI**

Süheyla GÖNÜL
BİYOİSTATİSTİK VE TIP BİLİŞİMİ ANABİLİM DALI
YÜKSEK LİSANS TEZİ

DANIŞMAN
Doç. Dr. Sıddık KESKİN

VAN- 2010

T. C.
YÜZÜNCÜ YIL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

**DOĞRUSAL OLMAYAN KANONİK KORELASYON ANALİZİ
VE SAĞLIK ALANINDA UYGULAMASI**

Süheyla GÖNÜL
BİYOİSTATİSTİK VE TIP BİLİŞİMİ ANA BİLİM DALI
YÜKSEK LİSANS TEZİ

İmza
Unvanı, Adı Soyadı
Jüri Başkanı

İmza
Unvanı, Adı Soyadı
Üye

İmza
Unvanı, Adı Soyadı
Üye

TEZ KABUL TARİHİ

TEŞEKKÜR

Bu konuyu yüksek lisans tezi olarak öneren ve bu çalışmanın araştırma sürecinden yazım aşamasına kadar olan süreçte bilimsel desteğini gördüğüm önemli deneyim ve bilgilerinden faydalandığım hocam Sayın Doç. Dr. Sıddık KESKİN'e teşekkürlerimi sunarım.

İÇİNDEKİLER

Kabul ve Onay	II
Teşekkür.....	III
İçindekiler	IV
Simgeler ve Kısaltmalar	VI
Tablolar	VII
Şekiller	VIII
1. GİRİŞ	1
2. GENEL BİLGİLER	5
3. GEREÇ VE YÖNTEM	9
3.1. Gereç	9
3.2. Yöntem	10
3.2.1. DOKKA’ da yaygın kullanılan istatistikler ve grafikler	11
3.2.2. DOKKA’nın varsayımları (ön şartları)	11
3.2.3. DOKKA için temel eşitlikler	12
3.2.4. Optimal ölçekleme seviyeleri	13
3.2.5. Optimizasyon.....	14
3.2.6. Maksimum rank sayısı.....	17
3.2.7. Marjinal frekanslar.....	18
3.2.8. Her setteki kayıp.....	18
3.2.9. Özdeğer.....	18
3.2.10. Çoklu uyum.....	18
3.2.11. Tekli uyum.....	18
3.2.12. Tekli kayıp.....	18
3.2.13. Bileşen yükleri.....	19

3.2.14. Kategori ölçümleri.....	19
3.2.15. Kategori sentroidleri	19
3.2.16. Planlanmış kategori sentroidleri	19
4. BULGULAR ve TARTIŞMA	20
4.1. Birinci Uygulama	20
4.2. İkinci Uygulama	28
4.3. Üçüncü Uygulama	34
4.4. Dördüncü Uygulama	41
5. SONUÇ	49
ÖZET	51
SUMMARY	52
KAYNAKLAR	53

SİMGELER ve KISALTMALAR

ALS	: Dalgalı En Küçük Kareler Algoritması
CCA	: Kanonik Korelasyon Analizi
DM	: Diabetes Mellitus
DOKKA	: Doğrusal olmayan Kanonik Korelasyon Analizi
İTÖ	: İstatistik Tutum Ölçeği
NLCCA	: Nonlinear Canonical Corelation Analysis
OVERALS	: Doğrusal Olmayan Kanonik Korelasyon Analizi
VKİ	: Vücut Kitle İndeksi

TABLÖLAR

Tablo 1. Ele alınan değişkenler ve özellikleri.....	9
Tablo 2. İncelenen özelliklere ait sayı ve yüzdeler.....	20
Tablo 3. Boyutlar için özdeğerler ve kayıp değerleri (1).....	21
Tablo 4. Değişkenlerin boyutlardaki ağırlıkları (1).....	22
Tablo 5. Değişkenlere ait bileşen yükleri (1).....	23
Tablo 6. Boyutlara ait uyum değeri (1).....	25
Tablo 7. Değişkenlere ait kategori koordinatları (1).....	26
Tablo 8. Boyutlar için özdeğerler ve kayıp değerleri (2).....	29
Tablo 9. Değişkenlerin boyutlardaki ağırlıkları (2).....	29
Tablo 10. Değişkenlere ait bileşen yükleri (2).....	30
Tablo 11. Boyutlara ait uyum değeri (2).....	31
Tablo 12. Değişkenlere ait kategori koordinatları (2).....	32
Tablo 13. Boyutlar için özdeğerler ve kayıp değerleri (3).....	35
Tablo 14. Değişkenlerin boyutlardaki ağırlıkları (3).....	35
Tablo 15. Değişkenlere ait bileşen yükleri (3).....	36
Tablo 16. Boyutlara ait uyum değeri (3).....	37
Tablo 17. Değişkenlere ait kategori koordinatları(3).....	39
Tablo 18. Boyutlar için özdeğerler ve kayıp değerleri (4).....	41
Tablo 19. Değişkenlerin boyutlardaki ağırlıkları (4).....	42
Tablo 20. Değişkenlere ait bileşen yükleri (4).....	43
Tablo 21. Boyutlara ait uyum değeri (4).....	44
Tablo 22. Değişkenlere ait kategori koordinatları (4).....	46

ŞEKİLLER

Şekil 1. Değişkenlere ait bileşen yükleri (1).....	24
Şekil 2. Değişkenlere ait çoklu kategori koordinatları (1).....	27
Şekil 3. Değişkenlerin kategorilerine ait sentroid koordinatları (1).....	28
Şekil 4. Değişkenlere ait bileşen yükleri (2).....	30
Şekil 5. Değişkenlere ait çoklu kategori koordinatları (2).....	33
Şekil 6. Değişkenlerin kategorilerine ait sentroid koordinatları (2).....	34
Şekil 7. Değişkenlere ait bileşen yükleri (3).....	37
Şekil 8. Değişkenlere ait çoklu kategori koordinatları (3).....	40
Şekil 9. Değişkenlerin kategorilerine ait sentroid koordinatları (3).....	41
Şekil 10. Değişkenlere ait bileşen yükleri (4).....	44
Şekil 11. Değişkenlere ait çoklu kategori koordinatları (4).....	47
Şekil 13. Değişkenlerin kategorilerine ait sentroid koordinatları (4).....	48

1. GİRİŞ

Artan dünya nüfusu ile birlikte, insan ihtiyaçlarını karşılamaya ve yaşam zorluklarını gidererek, yaşamın kalitesini yükseltmeye yönelik yapılan çalışmalar, beraberinde farklı bilim dallarının gelişmesine neden olmuştur. Bu bilim dalları temelde; doğadaki olaylara ilişkin oluşum mekanizmalarını belirleme ve buna bağlı olarak da bu olayların seyrini insan yaşam kalitesini yükseltmeye yönelik değiştirebilme eksenindedir. Bu bilim dallarının gelişim süreci; beraberinde bu süreçte kullanılan yöntemlerle ilgili sorunları da getirmiştir. Doğadaki olayların bağlı olduğu yasaları incelemeye yönelik gelişmiş olan pozitif bilimlerde amaç, özetle; üzerinde durulan konu ile ilgili olarak, sınırlı sayıdaki gözlemlerden hareketle o konuyu açıklayabilecek akla yatkın yöntem ve teorileri, bunları destekleyen kanıtlar yardımıyla geliştirmek ve sonraki araştırmalara ışık tutabilmektir (Keskin, 2002). Şüphesiz ki doğadaki olaylar, birbirleri ile ilişkili olan çok sayıdaki değişkenin veya faktörün ayrı ayrı veya birlikte etkileşimleri ile ortaya çıkan karmaşık yapıdaki bir durumdur. Bu yapı içerisinde; ilgilenilen olaya etkili olduğu düşünülen değişkenlerin birbirleri ile olan ilişkilerini çözerek, ilgilenilen olayın oluşum mekanizmasını basit ve anlaşılır yapıda açıklayabilmek oldukça zordur. Zira sınırlı zaman içerisinde ve sınırlı sayıdaki veri ile bu yapıyı doğru ve güvenilir bir şekilde ortaya koyabilmek için olayın durumuna ilişkin istatistik yöntemlerin doğru kullanılması gerekmektedir.

İlgilenilen değişken ile buna etkili olabileceği düşünülen diğer değişken veya değişkenler arasındaki ilişkiler; genel olarak doğrusal veya doğrusal olmayan ilişkiler olmak üzere iki başlık altında incelenebilir. Genel olarak doğrusal ilişkileri belirlemede; regresyon ve korelasyon analizi yöntemlerinin en sık kullanılan yöntemler olduğu söylenebilir.

Özellikler arasındaki doğrusal ilişkiyi belirlemede ilk akla gelen yöntem, bu özellikleri ikiye ikiye ele alarak, bunlar arasında Pearson korelasyon katsayısını hesaplamaktır. Pearson korelasyon katsayısı, bazı varsayımlar ya da ön şartlar yerine geldiğinde, sürekli değişkenler arasındaki doğrusal ilişkinin yönünün ve derecesinin belirlenmesinde en yaygın kullanılan ölçüdür (Keskin ve Özsoy 2004). Ancak bu

katsayının hesaplanması için gerekli olan ön şartlar yerine gelmemişse, Spearman rank korelasyonu ve Kendal tau korelasyonu gibi parametrik olmayan yöntemler kullanılabilir. Bazı durumlarda ele alınan özelliklerden birisi bağımlı, diğeri de bağımsız değişken olabilir. Bu durumda, iki değişken arasında hesaplanan korelasyon katsayısı, bu değişkenler arasındaki sebep - sonuç ilişkisinin bir ölçüsü olarak da değerlendirilebilir (Keskin ve Özsoy 2004). Bağımlı değişken bir adet, ancak bağımsız değişkenler birden fazla ise bunlar arasındaki ilişkiyi belirlemede çoklu korelasyon katsayısı (multiple correlation coefficient, determination coefficient) kullanılır. Bazı durumlarda, iki değişken arasındaki ilişkinin yönüne ve derecesine, bunlardan biri veya ikisi ile ilişkili olan başka değişken veya değişkenler etki ediyor olabilir. Bu gibi durumlarda, adı geçen iki değişken arasındaki ilişkiye etki eden değişken veya değişkenlerin etkilerinin giderilmesi işlemine dayalı olarak hesaplanan kısmi korelasyon katsayısının (partial correlation coefficient) hesaplanması daha uygundur. Hem bağımlı değişken, hem de bağımsız değişken birden fazla olduğu durumda, yani iki değişken seti (kümesi) arasındaki ilişkinin belirlenmesi istendiğinde, yukarıda belirtilen korelasyon katsayılarından hiç birisinin kullanımı uygun olmaz. Bunların yerine, değişken setleri ya da kümelerini, bu setlerde yer alan değişkenlerin doğrusal (linear) bileşenlerinden oluşan kanonik değişkenlere dönüştürerek, bu kanonik değişkenler arasındaki ilişki yapısını bulma temeline dayalı kanonik korelasyon kullanılır (Gürbüz, 1989).

Değişkenler arasındaki ilişkinin yönünü ve derecesini belirlemede değişkenlerin biri veya bir kısmı üzerine diğer değişkenlerin etkisinin belirlenmek istendiği çalışmalarda genelde regresyon analizi yöntemleri kullanılmaktadır. Regresyon analizi kısaca, araştırmacının üzerinde durduğu özellik veya özelliklerle, bu özelliğe etkili olabileceği varsayılan diğer özellikler arasındaki (doğrusal veya doğrusal olmayan) ilişkileri belirlemek üzere eşitliklerin bulunması ve bu eşitlikteki katsayıların yorumlanması işlemlerini kapsar (Keskin ve ark., 2007). Regresyon analizinde, ilgilenilen değişken; bağımlı değişken (Y), tahmin edilen değişken veya cevap değişkeni olarak adlandırılırken, bu değişkenle ilişkili olabileceği düşünülen değişken veya değişkenler ise bağımsız değişken (X), tahmin edici değişken veya açıklayıcı değişken olarak adlandırılır. Değişkenler arasındaki doğrusal ilişkileri belirlemek üzere

kurulacak olan regresyon eşitlikleri; bir cevap değişkeni ile bir açıklayıcı değişkeni içeriyorsa basit regresyon, bir cevap değişkeni ile birden çok açıklayıcı değişkeni içeriyorsa da, çoklu regresyon (multiple regression) olarak adlandırılır (Weisberg, 2005). Bazı durumlarda, birden çok Y değişkeni ile birden çok X değişkeni arasındaki ilişki yapısı da incelenmek istenebilir. Bu gibi durumlarda; ilk olarak, her bir Y değişkeni için ilgili X değişkenleri ile çoklu regresyon eşitliklerinin bulunması akla gelebilir. Ancak, bu durumda; ilgili Y değişkenleri arasında korelasyonların olmadığı, diğer bir ifade ile Y değişkenlerinin birbirinden bağımsız olduğu varsayılmaktadır. Genellikle, yapılması planlanan bilimsel çalışmalarda; Y değişkeni olarak düşünülen değişkenlerin, birbirinden bağımsız olmadığı, yani birbirleri ile ilişkili olduğu düşünülürse, her bir Y değişkeni için ayrı ayrı X değişken kümesi ile ilişkisinin bulunması yaklaşımı, değişkenler arasındaki gerçek ilişki yapısını belirlemede yanıltıcı sonuçlar verebilir. Bu gibi durumlarda; (cevap değişkeni ve açıklayıcı değişkenlerin birden çok olduğu durumda) kanonik korelasyon analizi ve çok değişkenli regresyon analizi yöntemlerinden yararlanılması daha uygun bir yaklaşım olabilir. Kanonik korelasyon analizi; genellikle X ve Y olarak adlandırılan iki değişken kümesi arasındaki ilişkiyi maksimum yapacak şekilde, bu kümelerde yer alan değişkenlerin doğrusal kombinasyonlarından oluşan yeni değişkenlerin bulunması ve elde edilen sonuçların yorumlanması işlemlerini kapsayan çok değişkenli analizi yöntemidir. Çok değişkenli regresyon analizi ise; X ve Y değişken kümeleri arasındaki çoklu doğrusal regresyon analizini, çok değişkenli yaklaşımla çözümlemeye yönelik geliştirilmiş bir analiz yöntemidir (Keskin ve ark.,2007).

Herhangi bir zorunluluk olmamasına rağmen, değişken kümelerinden birisi bağımlı, diğeri ise bağımsız değişken kümesi olarak düşünülebilir. Kanonik korelasyon analizi, yalnızca iki değişken kümesi arasındaki doğrusal ilişki yapısını inceler. Bunun yanı sıra, kanonik korelasyonda; değişkenlerin normal dağılım göstermesi, sürekli olması ve aralarındaki ilişkilerin doğrusal olması gibi ön şartlar bulunmaktadır. Uygulamada araştırmacılar; 2’den fazla değişken kümesi arasındaki ilişki yapısını da incelemek isteyebilir. Bu durumda, değişken kümelerini ikili ikili ele alıp, kanonik korelasyon ile incelemek akla gelse de bu durum, değişken kümeleri arasındaki ilişki yapısının bütünlüğünü bozacağından dolayı doğru bir yaklaşım olmayabilir. Ayrıca

kullanılan veriler, kanonik korelasyon analizinin kullanılabilmesi için yukarıda belirtilen ön şartları da sağlamayabilir. Bu durumda verilerin analizi için kullanılabilecek en iyi yöntem, Doğrusal Olmayan Kanonik Korelasyon Analizi (Nonlinear Canonic Correlation Analysis) dir. Değişkenlerin normallik varsayımını sağlamadığı, ya da değişken kümelerinin kesikli, kategorik ve sürekli değişkenlerden meydana geldiği bu gibi durumlarda, kanonik korelasyon analizinin etkinliği azalır. Bunun yerine OVERALS olarak da bilinen Doğrusal Olmayan Kanonik Korelasyon Analizinin kullanılması önerilebilir. Özetle, değişken kümesi sayısının 2 veya 2' den fazla ve değişkenlerin sürekli, sıralı veya kategorik olması durumunda; bu değişken kümeleri arasındaki doğrusal ve doğrusal olmayan ilişkiyi incelemede kullanılan yöntem Doğrusal Olmayan Kanonik Korelasyon Analizi (DOKKA) olarak bilinir. DOKKA değişken kümelerini oluşturan değişkenlerin, sürekli ve aralarındaki ilişkinin de doğrusal olması gerekliliği varsayımını içermez. Dolayısıyla, değişken tipi ve aralarındaki ilişki bakımından kanonik korelasyon analizine göre oldukça esnektir.

DOKKA, ilk olarak Gifi (1990), daha sonra da Van der Burg ve ark. (1984) tarafından önerilen çok değişkenli bir analiz yöntemidir. Yöntemin yaklaşık yirmi beş yıllık bir geçmişi olmasına rağmen, tanıtım eksikliği ve kullanımında bilgisayara gereksinim duyulması gibi nedenlerden dolayı uygulamada yaygın bir kullanım alanı bulamamıştır. DOKKA nominal (sınıflayıcı), ordinal (sıralayıcı), nümerik (sayısal) ya da bunların herhangi bir kombinasyonundan oluşan değişkenlerden meydana gelen iki veya daha fazla değişken seti arasındaki ilişkilerin incelenmesi ve benzerliklerin ortaya konulması bakımından oldukça kullanışlı bir yöntemdir. Ayrıca bu analiz tekniği çoklu doğrusal bağlantı nedeniyle incelenemeyen ilişkilerin analizine olanak sağlamakla birlikte kanonik korelasyon analizinde olduğu gibi boyut indirgeme yapabilmektedir. DOKKA iki değişken seti (kümesi) arasındaki ilişkiyi inceleyen kanonik korelasyon analizinin (k değişken seti için) genelleştirilmiş hali olarak düşünülebilir.

Bu çalışmada, DOKKA tanıtılarak, teorik altyapısı ile birlikte işlem aşamaları ayrıntılı olarak incelenecek ve sağlık alanında gerçek veri kümesi üzerinde uygulama yapılarak elde edilen sonuçlar yorumlanacaktır.

2. GENEL BİLGİLER

Glop (1985), ailelerin seyahat yapmalarında etkili olan haneye ve kişiye ait özellikleri dikkate alarak bunlar arasındaki ilişkileri standart CCA ve NLCCA ile incelemiştir. Sonuçta; yaş, yaşam biçimi ve gelir değişkenlerinin etkili değişkenler olduğunu vurgulamıştır.

Van der Burg ve De Leeuw (1988), genelleştirilmiş doğrusal olmayan kanonik korelasyon analizinde kategori ölçeklerinin, özdeğerlerin ortalamasını ve standart hatalarını; Bootstrap ve Jackknife örneklem yöntemini kullanarak tahmin etmişlerdir. Sonuçta genel olarak Bootstrap metodunun Jackknife metodundan daha iyi olduğunu ve geniş örneklerde daha iyi güven aralığı sonucunu verdiğini vurgulamışlardır. Bununla birlikte, geniş örneklerde hesaplanan özdeğerlerin daha istikrarlı olduğunu belirtmişlerdir.

Michailidis ve de Leeuw (1996), Doğrusal Olmayan Kanonik Korelasyon Analizi için kayıp fonksiyonu ve normalizasyon kısıtlamaları ile ilgili çalışmalar yapmışlardır.

Bohacek ve Jonckheere (1998), NLCCA'nın standart kanonik korelasyon analizinin genelleştirilmiş hali olduğunu vurgulayarak, doğrusal olmayan işlemlerde diğer bir ifade ile değişken kümeleri arasında doğrusal ilişkinin olmadığı durumlarda doğrusal olmayan kanonik korelasyon katsayılarının, standart kanonik korelasyon analizinden elde edilen katsayılara göre daha yüksek bulunacağını vurgulamışlardır.

Türe ve ark. (2000), Kardiyolojik verilerin DOKKA ile incelenmesine yönelik yaptıkları çalışmada, Trakya Üniversitesi Tıp Fakültesi Kardiyoloji kliniğinde 2000 yılına kadar koroner anjiyografisi yapılan 526 hastanın koroner arter çapları (1. set), koroner arter çaplarında lezyon olup olmadığı (2. set), cinsiyet, diabetes mellitus (DM), sigara kullanımı ve vücut kitle indeksi (VKİ) (3. set) değişkenleri ele alınarak bu üç setteki değişkenler arasındaki ilişkileri DOKKA ile incelemişler ve analiz sonucunda iki boyutlu çözümlemede birinci boyutun değişkenler arasındaki ilişkinin %54.7'sini, ikinci boyutun ise %44.3'ünü açıkladığını belirtmişlerdir.

Hsieh (2000), NLCCA'ni üçlü ileri beslenimli sinir ağıları modelinde kullanmıştır.

Bayram ve Ertaş (2001), Bursa hane halkı tüketim harcamaları verilerini kullanarak tüketici harcama düzeni ve biçimini belirlemeye yönelik araştırmalarında doğrusal olmayan temel bileşenler analizi (PRINCALS) ve OVERALS olarak ta bilinen doğrusal olmayan kanonik korelasyon analizini kullanmışlardır. Çalışmada önce 953 hane halkı için tüm değişkenleri bireysel olarak analiz etmek için PRINCALS, sonrasında da değişkenleri kümeler halinde ele alarak OVERALS analizi uygulamış ve sonuçta PRINCALS ve DOKKA'nın değişik ölçüm düzeylerine sahip çoklu bağıntılı değişkenler arasındaki ilişkileri incelemede yararlı olacağını vurgulamışlardır.

Hsieh (2001), yapmış olduğu çalışmada sinir ağıları modelinde DOKKA'nın kullanılabilirliğini incelemiştir. Daha önce yapılan çalışmalarda sinir ağıları yaklaşımında sentetik verilerle inceleme yapıldığını vurgulayarak, bu çalışmada gerçek verilerle DOKKA'nın kullanılabilirliğini incelemiştir.

Süt (2001), yaptığı çalışmada birinci sette cinsiyet, medeni durum, ünvan, idari görev ve görev yeri değişkenlerini, ikinci sette sigara ve alkol kullanımı değişkenlerini ve üçüncü sette de güvensizlik, endişe, mutluluk ve hassasiyet değişkenlerini alarak bunlar arasındaki ilişkileri DOKKA ile belirlemiştir. Sonuçta, cinsiyet ve alkol kullanımı değişkenleri ile hassasiyet değişkeni arasında, ünvan, görev yeri ve mutluluk değişkenleri ile sigara kullanımı değişkeni arasında ilişki olduğunu belirlemiştir.

Kumar ve ark. (2002), yapmış oldukları çalışmada radyal tabanlı fonksiyonlarda NLCCA'nın kullanılabilirliğini önermişlerdir ve çalışmalarında daha çok metottaki hataların belirlenmesi ve teşhisi üzerine yoğunlaşmışlardır.

Wu ve Hsieh (2002), sinir ağıları yaklaşımı ile aylık elde edilen Tropikal Pasifik'teki deniz yüzey sıcaklığı ve yüzeydeki rüzgar etkisini belirlemek amacıyla DOKKA'yı kullanmışlardır. DOKKA sonuçları CCA ile karşılaştırıldığında DOKKA'nın iki set arasındaki varyasyonu daha iyi açıklayabildiğini ve daha yüksek kanonik korelasyonlar elde edildiğini vurgulamışlardır.

Kruger ve ark. (2004), yeni bir NLCCA yaklaşımı algoritmasını tanıtmışlardır. Bu yaklaşımda öncelikle CCA'ne benzer şekilde her iki değişken setindeki değişkenlerin doğrusal kombinasyonlarını almışlar ve bu kombinasyonların doğrusal olmayan kanonik değişkenleri elde etmek üzere doğrusal olmayan transformasyona tabi tutmuşlardır. Doğrusal olmayan transformasyonlar için karesel ve kübik polinomları radyal tabanlı yapay sinir ağları fonksiyonları ile birlikte kullanmışlardır.

Yeung ve Tsang (2005), yapmış oldukları çalışmada fonksiyonel MR görüntüsü ile kan oksijen seviyesine ait çeşitli değişkenler arasındaki ilişkileri, DOKKA ile incelemişler ve elde edilen sonuçları 2 ve 3 boyutlu uzayda karşılaştırmışlardır.

Saraçlı ve Saraçlı (2006), Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi öğrencilerinin demografik özellikleri ile üniversite sorunları arasındaki ilişkiyi DOKKA ile incelemişlerdir. Yapmış oldukları bu çalışmada üniversitenin farklı bölümlerinde 2003-2004 eğitim-öğretim yılında okuyan 664 öğrenciye anket uygulamışlar ve sonuçta; incelenen değişken kümeleri arasındaki uyumu ifade eden özdeğerleri; birinci boyut için 0.827, ikinci boyut için de 0.762 olarak bulmuşlardır. Bu değerlerin karekökü olan kanonik korelasyon katsayıları ise birinci boyut için 0.909 ve ikinci boyut için 0.873 olarak hesaplanmıştır.

Theodosiu ve ark. (2007), yaptıkları çalışmada biyomedikal veri Setlerinde DOKKA'nın kullanılabilirliğini inceleyerek, bu analizin diskriminant analizi, naive Bayes ve destek vektör makineleri gibi kümeleme ve sınıflandırma süreçlerinde ilk aşama olarak kullanılabileceğini vurgulamışlardır.

Girginer ve ark. (2007), yapmış oldukları bir çalışmada, istatistiğe yönelik tutumlarda üniversite öğrencileri arasındaki bireysel farklılıkları DOKKA ile incelemişlerdir. Bu çalışma sürecinde öğrencilerin istatistiğe yönelik tutumlarını ölçmek için İstatistik Tutum Ölçeği (İTÖ) ve araştırmacılar tarafından hazırlanan demografik bilgi formu kullanılmıştır. Araştırmacılar bu çalışma için 2005–2006 eğitim öğretim yılında, Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi'nin İşletme, İktisat ve Maliye bölümlerinde okuyan 257 öğrenciye olumlu ve olumsuz yargılardan meydana gelen 24 maddelik İTÖ ile oluşturulan anketleri uygulamışlardır.

Sonuçta, öğrencilerin istatistik dersine karşı tutumlarını olumsuz yönde etkileyen en önemli nedenlerin İstatistik 1 ve İstatistik 2 dersindeki başarı notları, bu derslerin tekrar sayısı ve tekrar ediliş nedenleri olarak belirlenmiş ayrıca bölüm, cinsiyet, barınılan yer ve mezun olunan lise türü değişkenleri ile istatistik dersine yönelik tutumlar arasında anlamlı bir ilişki olmadığı tespit edilmiştir.

Cannon ve ark. (2008), dönemsel iklim verilerinin tahmininde robust NLCCA'ni kullanmışlardır. NLCCA'nin çok katmanlı sinir ağıları çerçevesinde kanonik korelasyon analizine göre daha esnek bir model olduğunu vurgulamışlardır. NLCCA'ni robust yapabilmek için modeldeki robust olmayan maliyet fonksiyonlarını robust maliyet fonksiyonlarıyla yer değiştirmişlerdir.

Frie ve Janssen (2009), sosyal eşitsizlik, sağlık ve yaşam üzerine yapmış oldukları çalışmada; 30 ile 59 yaş arası, rasgele alınan 695 katılımcıyla yapılan görüşmeler sonucunda; elde edilen sosyo ekonomik durum, yaşam koşulları ve tarzları, sağlıkla ilgili davranışlar ve sağlık değişkenleri ile ilgili verileri DOKKA ile analiz etmişler ve sonuçları yorumlamışlardır. Sonuçta, setler arasındaki uyumun 1.444 olduğu ve özdeğerlerin varyansın % 50 sini açıkladığı vurgulanmış ve değişkenler arasındaki ilişkiler üç boyutlu uzayda görüntülenmiştir.

Cannon (2006), yapmış olduğu bir çalışmada Lorenz eşitlik sistemlerinde elde ettiği çözümlerde Doğrusal Olmayan Temel Bileşenler Analizi ve Doğrusal Olmayan Kanonik Korelasyon Analizinin performanslarını değerlendirmiştir.

Yazıcı ve ark. (2010), yapmış oldukları çalışmada, K adet kategorik değişken kümesi arasındaki ilişkileri ve veri kümelerinin yapısal benzerliklerini DOKKA ile inceleyerek, tıp alanında yapılan çalışmalardan elde edilen geniş kapsamlı veri setlerinin analizinde yöntemin uygulama ve yorum bakımından kullanılabilirliğini araştırmışlardır.

3. GEREÇ VE YÖNTEM

3.1. Gereç

Bu çalışmada uygulama materyali olarak; Boysan ve ark. (2009), tarafından Yüzüncü Yıl Üniversitesi'nde 548 öğrenci ile yapılan anket çalışması sonucunda elde edilen verilerin bir kısmı kullanılmıştır. Çalışmada ele alınan değişkenler ve özellikleri Tablo 1'de verilmiştir. Verilerin analizi için SPSS (ver: 13.0) istatistik paket programı kullanılmıştır (Anonymous, 2007).

Tablo 1. Ele alınan değişkenler ve özellikleri

Değişken adı	Kategoriler	Tür
Cinsiyet	(1) Kız (2) Erkek	Tekli Nominal
Medeni hal	(1) Evli (2) Bekar	Tekli Nominal
Yaş Grubu	(1) 17- 19 (2) 20- 25 (3) 26- 30 (4) 31- 35	Ordinal
Ruhsal hal	(1) R.Var (2) R.Yok	Tekli Nominal
Adres	(1) Aile Yanı (2) Yurt (3) Pansiyon (4) Evde tek (5) Evde ark.	Çoklu Nominal
İkametgah	(1) B.Şehir (2) Şehir (3) Kasaba (4) Köy	Çoklu Nominal
Ekonomi	(1) Düşük (2) Orta (3) Yüksek	Ordinal
Kardeş sayısı grubu	(1) 0- 5 (2) 6- 10 (3) 11- 15	Ordinal
İntihar	(1) İ.Evet (2) İ.Hayır	Tekli Nominal
Cinsel travma	(1) C.Var (2) C.Yok	Tekli Nominal
Fiziksel travma	(1) F.Var (2) F.Yok	Tekli Nominal
Cinsel+Fiziksel trav.	(1) C+F.Var (2) C+F.Yok	Tekli Nominal

3.2. Yöntem

Günümüzde çok değişkenli tekniklerden doğrusal olmayan transformasyonun kullanımı artmıştır. Bunun nedeni ise optimal ölçekleme analizinde istenilen kriterleri maksimum yapacak şekilde değişkenlerin doğrusal olmayan transformasyonlarını içerdiği için transformasyonlarda ölçümlerde karşılaşılabilecek zorlukların sınırlanabilmesi olarak gösterilebilir. Yani optimal ölçeklemenin kullanılış amacı değişkenlerin ölçüm düzeyleri tarafından konan kısıtlamaları dikkate alarak p sayıda ölçek oluşturmaya çalışmaktadır denilebilir. Böylelikle ölçümlerdeki sınırlamalar ile CCA kombinasyonu optimal ölçeklemeyi verir. Optimal ölçekleme DOKKA'yı da içerir. Orijinal değişkenlerin yerine optimal olarak ölçeklenmiş olan yeni değişkenler kullanılır. Böylelikle geometrik olarak bu orijinal değişkenlerin yerine yeni ölçeklenmiş değişkenler kullanılarak analizin başlangıcındaki kriter maksimize edilmeye çalışılır.

Her değişken için bir transformasyon kullanılır DOKKA içinde tekli transformasyon uygulanabildiği gibi çoklu transformasyonlar da uygulanabilir. Bu teknik OVERALS olarak adlandırılır. Optimal ölçekleme, hesaplamalarda, yineli dalgalı en küçük kareler yöntemini kullanmaktadır. Homojenlik analizi, doğrusal olmayan temel bileşenler analizi ve doğrusal olmayan kanonik korelasyon analizi bilinen üç temel optimal ölçekleme tekniğidir.

OVERALS olarak bilinen Doğrusal Olmayan Kanonik Korelasyon Analizi, optimal ölçekleme ile yapılan kategorik kanonik korelasyon analizi ile eşdeğerdir. Kanonik korelasyon analizi; genellikle X ve Y olarak adlandırılan iki değişken kümesi arasındaki ilişkiyi maksimum yapacak şekilde, bu kümelerde yer alan değişkenlerin doğrusal kombinasyonlarından oluşan yeni değişkenlerin bulunması ve elde edilen sonuçların yorumlanması işlemlerini kapsayan çok değişkenli analiz yöntemlerinden birisidir.

Optimal ölçekleme yaklaşımı, üç farklı açıdan standart kanonik korelasyonun genişletilmiş hali olarak düşünülebilir. Bunlardan birincisi, optimal ölçekleme, dolayısıyla DOKKA, ikiden fazla değişken kümesi arasındaki ilişkinin bulunmasını da sağlar. İkincisi, kümelerde yer alan değişkenler; istenilen ölçekte (nominal, ordinal ya

da nümerik) olabilir ve üçüncüsü doğrusal olmayan ilişkiler de analiz edilebilir. Bununla birlikte, Kanonik Korelasyondaki değişken kümeleri arasındaki korelasyonun maksimum yapılması yerine, DOKKA’ da deney üniteleri ya da nesnelere ait skor değerleri arasındaki uyum karşılaştırılır.

Optimal Ölçekleme ile Kategorik Kanonik Korelasyon Analizi veya DOKKA, değişken kümeleri arasındaki ilişkilerin istenilen boyutta grafiksel olarak gösterilmesine de olanak sağlamaktadır.

3.2.1. DOKKA’ da yaygın kullanılan istatistikler ve grafikler

Frekanslar, sentroidler yineleme (iterasyon), nesne skorları, kategorilerin rakamsal olarak belirlenmesi, ağırlıklar, bileşen yükleri, tekli ya da çoklu uyum (single and multiple fit), nesne skorlarının grafiği, kategori koordinatları, bileşen yüklerinin grafiği, kategori sentroidlerinin grafiği ve transformasyon grafikleridir.

3.2.2. DOKKA’ nın varsayımları (ön şartları)

DOKKA’ da veriler ile ilgili herhangi bir ön şart olmayıp, veriler nominal (isimsel), ordinal (sıralı) ya da nümerik (rakamsal) olabilir.

Kategorik yapıdaki değişkenlerin kodlanmasında genellikle 1’ den başlanıp birbirini takip eden ardışık rakamlar verilir.

Değişkenler, iki ya da daha çok kümeye ayrılabilir ve tekli nominal, çoklu nominal, sıralı ya da nümerik olarak ölçeklendirilebilir.

Elde edilebilecek maksimum boyut sayısı, değişkenlerin optimal ölçekleme sayısına bağlıdır. Eğer bütün değişkenler, tekli nominal, sıralı ya da nümerik olarak tanımlanmışsa, maksimum boyut sayısı “ya gözlem sayısı -1” den ya da toplam değişken sayısından daha azdır. Ancak yalnızca iki değişken kümesi varsa, maksimum boyut sayısı, standart kanonik korelasyon analizinde olduğu gibi küçük olan kümedeki değişken sayısı kadardır. Eğer bazı değişkenler çoklu nominal ise maksimum boyut

sayısı; çoklu nominal kategori sayısı ile çoklu olmayan nominal değişkenlerin toplamından çoklu nominal değişken sayısının çıkarılmasına eşittir. Örneğin analiz, 10 değişken içeriyor ve bunlardan biri 4 kategorili çoklu nominal ise maksimum boyut sayısı $(9 + 4 - 1 = 12)$ 12 dir.

Eğer her değişken kümesinde yalnızca bir değişken varsa, DOKKA optimal ölçekleme ile yapılan Temel Bileşenler Analizi ile eşdeğerdir. Eğer bu değişkenlerin her birisi çoklu nominal ise analiz, çoklu uyum analizi ile eşdeğerdir. Eğer iki değişken kümesi varsa ve bu kümelerden birisi yalnızca bir kategorik değişken içeriyorsa, analiz optimal ölçekleme ile yapılan kategorik regresyon analizi ile eşdeğerdir.

3.2.3. DOKKA için temel eşitlikler

n : Gözlem ya da deney ünitesi sayısı

m : Değişken sayısı

p : Boyut sayısı

K : Değişken kümesi sayısı

Değişken sayısı: $j = 1, \dots, m$ olmak üzere; j . değişken için k_j ; j . değişkenin geçerli kategori sayısı ya da uzaklık değeri; G_j de $n \times k_j$ boyutlu j . değişken için gösterge (indicator) matrisi olarak tanımlandığında;

$$g_{(j)ir} = \begin{cases} 1 & \text{dir. Eğer } i.\text{nesne } j.\text{değişkeninin } r.\text{ kategorisinde ise} \\ 0 & \text{dir. Eğer } i.\text{nesne } j.\text{değişkeninin } r.\text{ kategorisinde değilse} \end{cases}$$

$D_j : G_j$ 'nin sütun toplamlarından oluşan diyagonal matris olup; k . değişken kümesi ($k = 1, \dots, K$) için

$J(k) : k$. kümeye ait değişkenlerin indeks kümesi olup $j \in J(k)$ yazılabilir

mk : k . kümedeki değişken sayısı ($J(k)$ daki eleman sayısı)

M_j : Diyagonal elemanlar

$$m_{(j)ii} = \begin{cases} 1 & \text{dir. Eğer } \forall j \in J(k) \text{ için } i.\text{gözlem} [1, k_j] \text{ aralığında ise} \\ 0 & \text{dir. Eğer } \forall j \in J(k) \text{ için } i.\text{gözlem} [1, k_j] \text{ aralığında değilse} \end{cases}$$

olan $n \times n$ ikili diyagonal matris

X : $n \times p$ boyutlu nesne skorları matrisi

X_j : j . değişkene uyumlu nesne skorları ile düzeltilmiş $n \times p$ boyutlu yardımcı matris

Y_j : $k_j \times p$ boyutlu çoklu değişkenler için kategori değerleri

y_j : k_j için tekli değişkenlere ait kategori değerleri

a_j : p . boyut için tekli değişkenlere ait kategori değerleri

Q_k : $n \times m_k$ boyutlu matrisin $q_j = G_j y_j$ için k . kümedeki değişkenlerin ölçümleri

Y : Tekli ve çoklu kategorilerin değişkenler ve değişken kümeleri üzerinden toplamı

OVERALS fonksiyonunun amacı, belirli kısıtlamalar altında Y_j ($j=1, \dots, m$) ve X nesne skorlarının bulunmasıdır. Bu kısıtlamalar değişik şekillerde olabilmektedir.

I , $p \times p$ boyutlu birim matris ve $M^* = \sum_k M_k$ olmak üzere; $X' M^* X = k \cdot n \cdot I$

normalizasyon kısıtlaması altında, $\sigma(X, Y_j) = 1/$

$K \sum_k \text{tr}[(X - \sum_{j \in J(k)} G_j Y_j)' M_k (X - \sum_{j \in J(k)} G_j Y)]$ minimum yapılır.

3.2.4. Optimal ölçekleme seviyeleri

Çoklu nominal: $Y_j = Y_j$ (Yalnızca eşitlik kısıtlaması)

Tekli nominal: $Y_j = y_j a'_j$ (Eşitlik ve bir - rank kısıtlaması)

Tekli Ordinal: $Y_j = y_j a'_j$ ve $y_j \in C_j$ (Eşitlik, birim – ranklılık ve monotonluk (tekdüzelik) kısıtlaması) monotonluk kısıtlaması olan $y_j \in C_j$ ile y_j 'lerin; k_j vektörlerinin azalmayan elemanlarının dışbükey konide yer alması gerektiği kastedilmektedir.

Tekli nümerik: $Y_j = y_j a'_j$ ve $y_j \in L_j$ (Eşitlik, birim – ranklılık ve doğrusallık kısıtlaması). Doğrusallık kısıtlaması olan $y_j \in L_j$ ile y_j 'lerin; birbirini takip eden tamsayı değerler olan k_j 'lerin doğrusal transformasyonlarını içeren bütün k_j vektörlerinin alt uzayında yer alması gerektiği kastedilmektedir. Her değişken için bu seviyeler birbirinden bağımsız olarak seçilir. Bütün seçenekler için genel koşul ya da şart, eşit ölçeklemeyi alan göstergelerin eşit kategorili olmasıdır. Çoklu olmayan

seenekler iin genel art ya da koul ise $Y_j = y_j \mathbf{a}'_j$ olmasıdır. Burada Y_j , 1 ranklıdır. Tanımlama amaçlı olarak y_j daima normalizedir ve böylece $\mathbf{y}'_j \mathbf{D}_j \mathbf{y}_j = n_w$ dir.

3.2.5. Optimizasyon:

Optimizasyon aağıdaki iterasyon emasının icra edilmesi ile elde edilir.

- a) Balatma (Rasgele ya da i ie seim)
- b) Deėikenler ile deėiken kmelerinin (setlerin) baėlanması
- c) Diėer deėikenlerin etkilerinin veya katkılarının elemine edilmesi
- d) Kategori leklerinin gncellenmesi
- 3) Nesne skorlarının gncellenmesi
- f) Ortnormalizasyon
- g) Yakınsama testi 2-6. adımların tekrarı
- h) Rotasyon

a) Balatma

Rasgele (Random) seim : Nesne skorları olan $\mathbf{X}$ rasgele bir sayı ile balatılır. Sonra $\mathbf{X}$ normalize edilir. Bylece $\mathbf{u}'\mathbf{M}\mathbf{W}\mathbf{X} = 0$ ve $\mathbf{X}'\mathbf{M}\mathbf{X} = kn\mathbf{I}$ den $\tilde{\mathbf{X}}$ deėeri elde edilir.

oklu deėikenler iin balangı kategori lekleri 0 olarak atanır. Tekli deėikenler iin balangı kategori lekleri olan $\tilde{y}_j$ balangıtaki k_j olarak tanımlanır. Sonra birbirini takip eden tamsayılar. $\mathbf{u}'\mathbf{D}_j \tilde{y}_j = 0$ ve $\tilde{y}_j \mathbf{D}_j \tilde{y}_j = n$ eitliklerine gre normalize edilir ve balangı deėikeni 0 olarak aėırlıklandırılır.

İ ie (Nested) seim : İ ie yapıda yukarıdaki iterasyon eması iki kez icra edilir. Birinci dngde, btn tekli deėikenler geici olarak “tekli nmerik” deėiken olarak dikkate alınır. Bylece ikinci dngye hazırlanır. Btn ilikili nicelikler, birinci dng sonularından kopyalanır.

b) Değişkenler ile değişken kümelerinin (setlerin) bağlanması

$k=1,...,K$ ($j \in J_{(k)}$) için gelecek iki adım tekrar edilir. j . değişkenlerin güncellenmesi esnasında diğer değişkenlere ait geçerli olan değerinde sabit tutulur.

c) Diğer değişkenlerin etkilerinin veya katkılarının elemine edilmesi

k . setteki j . değişken için yardımcı matris; $V_{(k)j} = \sum_{j \in j(k)} G_j Y_j - G_j Y_j$ olup bu matris

k . değişken setindeki diğer değişkenlerin, katkılarını veya etkilerini toplar. Böylece diğer değişkenlerin katkıları nesne skorlarından elemine edilmiş olur. Bu kayıp değerinin X ve Y_j 'nin bir fonksiyonu olarak "Loss $\sigma(X; Y_j)$ " şeklinde yazılmasını sağlar. Buna göre; kayıp fonksiyonu $\sigma(X; Y_j) = \text{Sabit} + 1/K \text{ tr} [(X - V_{k(j)}) - G_j Y_j]' M_k [(X - V_{k(j)}) - G_j Y_j]$ olup, $\tilde{X}$ geçerli değerine göre sabitlenir.

$\tilde{Y}_j = G_j' M_k G_j)^{-1} G_j' M_k (\tilde{X} - V_{k(j)})$ eşitliğinden Y_j elde edilir. Bu eşitlik daha sonraki hesaplamalar için temel oluşturur. Aynı set içerisinde diğer değişken olan l ' değişkenine geçildiğinde; $V_{(k)l}$ matrisi yeniden hesaplanmaz ancak güncellenir.

d) Kategori ölçeklerinin güncellenmesi

Çoklu nominal değişkenler için yeni kategori $Y_j^+ = \tilde{Y}_j$ eşitliğine göre basitçe ölçeklendirilir.

Tekli değişkenler için Dalgali en küçük kareler algoritmasının (ALS) bir döngüsü $\tilde{Y}_j$ 'nin bir-ranklı hesaplanabilmesi için icra edilir. Döngü $a_j^+ = \tilde{Y}_j' D_j \tilde{y}_j$ ile $\tilde{y}_j$ 'nün önceki kategori ölçeklendirmesinden başlar. Eğer değişken nümerikse bu kullanılır. Değilse $y_j^* = \tilde{Y}_j a_j^+$ hesaplanır.

Geçerli ya da güncel değişken tekli nominal ise y_j^+ , y_j^* 'nin $y_j^* \leftarrow \text{WMON}(y_j^*)$ normalizasyonu ile elde edilir. Buradaki $\text{WMON}(\)$ notasyonu ağırlandırılmış monotonik regresyon sürecindeki ağırlıkları göstermekle birlikte bu y_j^* değerini monotonik artan yapar. Kullanılan ağırlıklar D_j nin diyagonal elemanlarıdır. Bundan alt algoritmalar kullanılır ve sonuç $y_j^+ = n_w^{1/2} y_j^* (y_j^* D_j y_j^*)^{-1/2}$ olarak normalize edilir ve sonuçta $\underline{Y}_j^+ = y_j^+ a_j^+$ olur.

e) Nesne skorlarının güncellenmesi

İlk olarak yardımcı skor değerleri matrisi olan W ; $W \leftarrow W + M_k \sum_{j \in J(k)} G_j \underline{Y}_j^+$ olarak hesaplanır ve M^* göre merkezileştirilir. Buradan $X^* = \{I - M^* u u' / u' M^* u\} W$ olur.

f) Ortonormalizasyon

Burada problem M^* dan ortonormal X^+ 'yı bulmaktır. Bu $M^{*-1} X^*$ 'e yakındır. İşlem M^* ağırlandırılmış en küçük kareler gibidir. . OVERALS' ta bu işlem $X^+ \leftarrow m^{1/2} M^{*-1/2} \text{PROCRU}(M^{*-1/2} X^*)$ eşitliği ile yapılır. Buradaki $\text{PROCRU}(\)$ notasyonu, Procrustes ortonormalizasyon sürecini göstermek için kullanılmaktadır. Eğer $M^{*-1/2} X^*$ girdi matrisinin tekil değer ayrışımı ve $K \Lambda' L' K' K = I$, $L' L = I$ ile gösterilirse Λ diyagonal elemanlar ise çıktı matrisi M^* ölçeğinde; $KL' = M^{*-1/2} X^* L \Lambda'^{-1} L'$ ortonormalliğini sağlar. Buradaki L ve Λ hesaplanması QL algoritmasını izleyen Householder transformasyonu ile üçlü diyagonalizasyon tabanlıdır.

g) Yakınsama testi

Birbirini takip eden $\text{tr } \Lambda^4$ iz değerleri arasındaki fark hesaplanır ve bu daha önce belirlenen ε gibi küçük bir değer olan yakınsama kriteri ile karşılaştırılır. Bu farklılık başlangıçta belirlenen ε değerini geçerse 2 - 6. adımlar tekrar edilir. Yakınsama sağlandığında uyum yetersizliği veya eksikliği (badness of fit) değeri verilir.

h) Rotasyon (Döndürme)

OVERALS kayıp fonksiyonu olan $\sigma(X; \underline{Y})$ X ve $\underline{Y}$ nin eşzamanlı rotasyonu altında değişmez. Çözüm ortalama gösterim operatörünün temel eksenleri ile ilişkilidir.

$$Q_* = 1/K \sum_k M_k Q_k \left(Q_k' M_k Q_k \right)^{-1} Q_k' M_k$$

Temel eksen konumunu elde edebilmek için rotasyon matrisinin bulunması yeterlidir. Bu işlem $M_*^{-1/2} X^*$ matrisinin çarpımından elde edilir. Bu diyagonal matris 6. adımdaki Procrustes ortonormalizasyonunda kullanılanla eşdeğerdir.

Bu bölümde $X^+ \underline{Y}^+$ matrislerine ve L matrisi ile a_j vektörlerine döndürme (rotasyon) yapılır. Döndürülmüş matris olan L , 6. adımda tanımlanan en son PROCRU işleminden alınmıştır.

3.2.6. Maksimum rank sayısı

Maksimum rank “ $\rho_{\max}$ ” herhangi bir veri seti için hesaplanabilen maksimum boyut sayısını belirtmektedir ve genellikle

$$\rho_{\max} = \begin{cases} \min\{(n-1), r_1, r_2\} & K = 2 \text{ ise} \\ \min\{(n-1), \max r_k\} & K > 2 \text{ ise} \end{cases} \quad \text{olarak hesaplanır.}$$

Buradaki r_k değeri, $r_k = \sum_{j \in JM(k)} k_j + m_{k1} - m_{k2}$ eşitliğindeki gibi tanımlanır. Buradaki

m_{k1} , k değişken setindeki kayıp gözlem içermeyen çoklu değişken sayısıdır. m_{k2} ise k değişken setindeki tekli değişken sayısıdır. $JM(k)$ indeks seti olup çoklu setlerdeki değişkenlerin sayısıdır. Aşağıdaki koşullardan herhangi birisi yerine gelmezse OVERALS durdurulur.

$$1) \ r_k < n_k - 1 \quad 2) \ n_k > 2 \quad 3) \ \sum_k r_k \leq \sum_k (n_k - 1) - (n_{\max} - 1)$$

buradaki n_k k . setteki kayıp gözlem içermeyen nesne sayısıdır. $n_{\max}$ ise bütün n_k lardaki nesne sayısını gösterir.

3.2.7. Marjinal frekanslar

Marjinal frekanslar her bir deęişken için kayıp gözlem içeren ve içermeyen deęerleri verir. D_j ve her $j \in J(k)$ için ve her $j \in J(k)$ için M_k matrisinin sütun toplamaları olarak hesaplanır.

3.2.8. Her setteki kayıp (Loss per set)

Kayıp, K adet $\sigma(X; \underline{Y})$ deęerinin boyut ve setler göre K kez bölünmesidir. Buradan her boyut için ortalama verilir.

3.2.9. Özdeęer

Özdeęerler her boyut için “ortalama+1” olarak elde edilir. $\rho - \sigma(X; \underline{Y})$ ve yakınsama saęlandığında bu deęerler Q_* ’ın özdeęerleri olur.

3.2.10. Çoklu uyum

Bütün deęişkenler (satırlar) için verilen sütun boyutları ile bu ölçü $\underline{Y}_j' D_j \underline{Y}_j$ matrisinin diyagonal elemanları olarak hesaplanır

3.2.11. Tekli uyum: (Single fit)

Karesel ağırlıklar tekil deęişkenler için tekli hesaplanır. $a_j' a_j$, ağırlıklara ait kareler toplamıdır.

3.2.12. Tekli kayıp

Tekli deęişken için çoklu uyum, tekli uyuma eşittir.

3.2.13. Bileşen yükleri (Tekli değişkenler için) (Component Loadings For Single Variables)

Yükler tekli değişkenler için rakamsal değerleri gösterimlerinin (izdüşümlerinin) q'_j X nesne uzayındaki uzunluklarıdır. Eğer kayıp gözlem yoksa yükler, temel bileşenler nesne skorları ile ölçülmüş değişkenler arasındaki korelasyonlardır.

3.2.14. Kategori ölçümleri (Category quantifications)

Tekli koordinatlar: Tekli değişkenler için $\underline{Y}_j = y_j a'_j$ eşitliğidir. Çoklu koordinatlar için ise bu değer daha önce tanımlanan $\tilde{Y}_j$ 'ye eşittir.

3.2.15. Kategori sentroidleri (Category centroids)

Paylaşılan aynı kategorideki nesnelerin sentroidleri $D_j^{-1} G'_j X$ dir. Bunların çoklu koordinatlarının eşit olma zorunluluğu yoktur.

3.2.16. Planlanmış kategori sentroidleri (Projected category centroids)

Tekli değişkenler için planlanmış kategori sentroidleri $y_j b'_j$ dir. Bu b_j yükü ile verilen grafikteki çizgi üzerindeki noktalardır ve bunlar ağırlıklandırılmış D_j kategori sentroidlerinin izdüşümünden elde edilir (Anonymous, 2007).

4. BULGULAR VE TARTIŞMA

Çalışmada, ele alınan değişken sayısı ve değişkenlerin bulunduğu kümeye bağlı olarak dört uygulama yapılmıştır. Birinci ve ikinci uygulamalarda değişkenler 3 set, üçüncü uygulamada değişkenler 4 set ve dördüncü uygulamada da değişkenler 2 set olarak alınmıştır.

4.1. Birinci Uygulama

Birinci uygulamada değişkenler 3 set olarak alınmıştır. Birinci sette; cinsiyet, yaş, ruhsal hal ve medeni hal değişkenleri, ikinci sette; adres, ikametgah, kardeş grubu ve ekonomi değişkenleri, üçüncü sette ise; intihar, cinsel travma, fiziksel travma ve cinsel+ fiziksel travma değişkenleri yer almıştır. Çalışmada ele alınan değişkenler için kategorilere göre sayı ve yüzde değerleri Tablo 2’ de verilmiştir. Tablo 2’de yer alan sayı ve yüzde değerleri diğer uygulamalar içinde geçerlidir.

Tablo 2. İncelenen özelliklere ait sayı ve yüzdeler

Cinsiyet	Sayı	Yüzde	Medeni hal	Sayı	Yüzde	Cin.Trv.	Sayı	Yüzde
Kız	183	33,4	Evli	18	3,3	Var	50	9,1
Erkek	365	66,6	Bekar	530	96,7	Yok	498	90,9
Fiz. Trv	Sayı	Yüzde	Cin+Fiz. Trv.	Sayı	Yüzde	İntihar	Sayı	Yüzde
Var	136	24,8	Var	42	7,7	Evet	37	6,8
Yok	412	75,2	Yok	506	92,3	Hayır	511	93,2
Kardeş	Sayı	Yüzde	Ekonomi	Sayı	Yüzde	Ruhsal hal	Sayı	Yüzde
0-5	323	58,9	Yüksek	130	23,7	1	48	8,8
6-10	199	36,3	Orta	376	68,6	2	500	91,2
11-15	26	4,7	Düşük	42	7,7	Adres	Sayı	Yüzde
Yaş	Sayı	Yüzde	İkamet	Sayı	Yüzde	AileYanı	135	24,6
17- 19	65	11,9	B.Şehir	135	24,6	Yurt	170	31,0
20- 25	452	82,5	Şehir	322	58,8	Pansiyon	17	3,1
26- 30	28	5,1	Kasaba	51	9,3	Evde tek	18	3,3
31- 35	3	0,5	Köy	40	7,3	EvdeArk.	208	38,0

Tablo 3. Boyutlar için özdeğerler ve kayıp değerleri

		Boyut		Toplam
		1	2	
Kayıp	Set 1	0.323	0.456	0.779
	Set 2	0.422	0.696	1.118
	Set 3	0.695	0.461	1.157
	Ortalama	0.480	0.538	1.018
Özdeğer		0.520	0.462	
Uyum				0.982
Açıklanan uyum %		52.95	47.04	

Tablo 3’de DOKKA sonucunda değişken setleri arasındaki optimal uyumu sağlayan boyutlara ilişkin kayıp ve uyum değerleri verilmiştir. Buna göre toplam kayıp, boyutlara ve değişken setlerine paylaştırılmıştır.

Bilindiği üzere, kayıp; her boyut ve değişken seti için optimal ölçekleme ile değişkenlerin ağırlıklı olarak kombine edilmesi sonucunda, nesne skorlarındaki açıklanamayan varyans oranını belirtir. Birinci ve ikinci boyuttaki ortalama kayıpların toplamı 1.018 olup, 2. boyuttaki kaybın (0.538) birinci boyuttaki kayıptan (0.480) daha yüksek olduğu görülmektedir. Tablo 3’ te verilen özdeğerler her boyut için 1-ortalama kayıp değeri olarak belirtilmektedir (1. boyut için $1-0.480=0.520$, 2. boyut için $1-0.538=0.462$).

Toplam uyum olan 0.982 değerine göre 1. boyutla gerçek uyumun %52.95 i açıklanabilmektedir ($0.520/0.982$). Benzer şekilde 2. boyut ise gerçek uyumun %47.04 ünü açıklayabilmektedir ($0.462/0.982$). Maksimum uyum değeri boyut sayısına eşit olup, bu durum tüm veri setleri ile boyutlar arasındaki mükemmel uyumu göstermektedir. Toplam ortalama kayıp ile toplam uyum, boyut sayısına eşittir ($1.018+0.982=2$).

Bilindiği üzere, 2 değişken kümesi arasındaki en önemli istatistik kanonik korelasyon olup, kanonik korelasyon da özdeğerlerle ilişkili olduğundan, genellikle DOKKA’da kanonik korelasyonlar yerine özdeğerler verilir. Ancak istenirse 2 değişken seti arasındaki kanonik korelasyonlar $r=2xE_d-1$ eşitliği ile hesaplanabilir. Bu eşitlik 2 den fazla setler arasındaki kanonik korelasyon için $r_d = ((KxE_d)-1)/(K-1)$ dir.

Burada d boyut sayısı, K ise set sayısıdır. Buna göre 1. boyut için kanonik korelasyon $((3 \times 0.520) - 1) / (2) = 0.28$ dir. İkinci boyut ise $((3 \times 0.462) - 1) / (2) = 0.193$ dir.

Bir diğer ilişki ölçüsü nesne skorları ile her setten hesaplanan doğrusal kombinasyonlar arasındaki çoklu korelasyonlardır. Bir sette çoklu nominal olarak tanımlanmış bir değişken yoksa bu ölçü setteki her değişkenin bileşen yükü ile ağırlık çarpımına bu çarpımların eklenmesi ve karekökünün alınması ile hesaplanır. Örneğin birinci setteki cinsiyet, yaş, ruhsal hal ve medeni hal için çoklu korelasyon (R) her bir boyut için 1-kayıp (kayıp) = R^2 dir.

Değişkenlerin boyutlardaki ağırlıkları ve bu değişkenlere ait bileşen yükleri sırasıyla tablo 4 ve 5'te verilmiştir.

Tablo 4. Değişkenlerin boyutlardaki ağırlıkları

Set		Boyut	
		1	2
1	Cinsiyet	0.726	0.094
	Medeni hal	0.040	-0.350
	Yaş Grubu	0.223	0.016
	Ruhsal hal	0.058	-0.642
2	Ekonomi	-0.247	0.150
	Kardeş sayısı grubu	0.291	0.110
3	İntihar	-0.342	0.514
	Cinsel Travma	0.212	0.005
	Fiziksel Travma	0.403	0.303
	Cinsel+Fiziksel Trav.	0.316	0.395

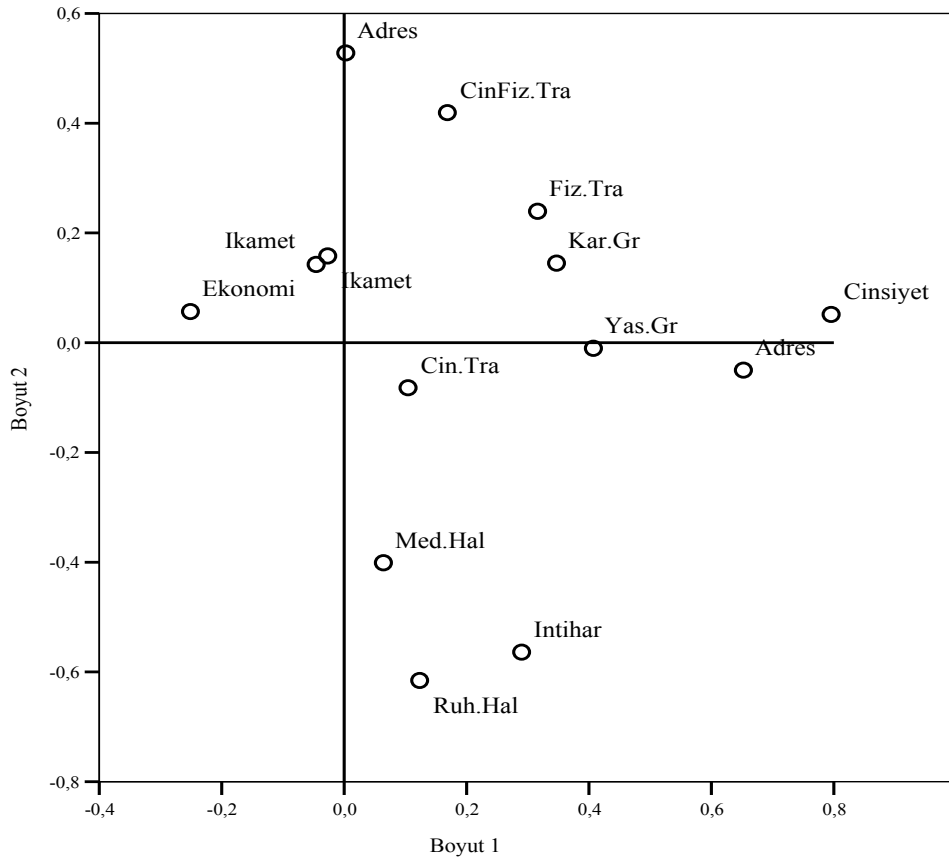
Tablo 4 de, değişkenlerin ağırlıkları yani boyutların oluşmasında değişkenlerin etkisi verilmiştir. Tablo 4 incelendiğinde; cinsiyet değişkeninin birinci boyuttaki ağırlığının 0.726, ikinci boyuttaki ağırlığının ise 0.094 olduğu görülmektedir. Buna dayanarak cinsiyet değişkeninin birinci boyut için daha iyi bir ayırım yaptığı söylenebilir. Aynı şekilde, ruhsal hal değişkeninin ağırlıkları incelenecek olursa ruhsal hal değişkeninin birinci boyut için oldukça düşük ancak ikinci boyutta oldukça kuvvetli bir ayırım yaptığı söylenebilir. Benzer şekilde, diğer değişkenler de incelenecek olursa yaş sayısı grubu, ekonomi, kardeş sayısı grubu, cinsel travma ve fiziksel travma değişkenlerinin birinci boyut için, medeni hal, intihar ve cinsel+fiziksel travma değişkenlerinin ise ikinci boyut için daha kuvvetli bir ayırım yaptıkları söylenebilir.

Tablo 5. Değişkenlere ait bileşen yükleri

				Boyut	
				1	2
1	Cinsiyet			0.789	0.055
	Medeni hal			0.057	-0.362
	Yaş Grubu			0.423	-0.025
	Ruhsal hal			0.123	-0.642
2	Adres	Boyut	1	0.653	-0.039
			2	0.014	0.498
	İkamet	Boyut	1	-0.044	0.150
			2	-0.021	0.172
	Ekonomi			-0.245	0.068
	Kardeş sayısı grubu			0.346	0.136
3	İntihar			-0.292	0.580
	Cinsel Travma			0.113	-0.091
	Fiziksel Travma			0.314	0.235
	Cinsel+Fiziksel Trav.			0.172	0.431

Bilindiği üzere, eğer veri setinde kayıp gözlem yoksa bileşen yükleri; nesne skorlarıyla, ölçeklendirilen değerler arasındaki Pearson korelasyon katsayıları olup, bu katsayılar Tablo 5’te verilmiştir.

Tablo 5 incelendiğinde birinci boyut ile en yüksek korelasyona sahip olan değişkenin 0.789 değeriyle cinsiyet değişkeni olduğu, bunu sırasıyla 0.653 ile adres, yaş grubu ve kardeş grubu değişkenlerinin izlediği görülmektedir. Benzer şekilde, ikinci boyutta ise en yüksek yük değeri -0.642 ile ruhsal hal değişkenine aittir. Bunu sırasıyla 0.580 ile intihar, 0.498 ile adres (2) ve 0.431 ile de cinsel+fiziksel travma değişkenlerinin izlediği görülmektedir. Değişkenlerin yük değerlerinin verilmiş olduğu Tablo 5, iki boyutlu uzayda grafiksel olarak Şekil 1’ deki gibi gösterilebilir.



Şekil 1. Değişkenlere ait bileşen yükleri

Şekil 1’ de incelenen değişkenler için bileşen yükleri görülmektedir. Her bir değişkenin orijinden uzaklaşması bu değişkenin önemliliği ile ilgilidir. Buna göre 1. boyut, daha çok cinsiyet, adres, yaş sayısı grubu ve ekonomi değişkenleri tarafından belirlenirken, 2. boyut ruhsal hal, adres ve intihar değişkenleri tarafından belirlenmektedir. Kanonik değişkenler çizilemezler ancak orijinden çıkan dikey ve yatay çizgilerle gösterilebilirler.

Tablo 6’da verilen tekli kayıp, bir setteki değişkenlerin ölçeklenmesi sonucu (tekli nominal, ordinal ya da nominal) kayıpları göstermektedir. Eğer tekli nominal büyük ise bu değişkenin çoklu nominal olarak alınması daha iyi olacaktır. Çalışmada cinsiyet değişkenine ait çoklu ve tekli uyum değerleri birbirine eşit çıkmıştır. Dolayısıyla, bu demektir ki, çoklu koordinatlar ağırlıklarla (weights) verilen doğrular üzerindedir.

Tablo 6. Boyutlara ait uyum değerleri

Set		Çoklu Uyum			Tekli Uyum			Tekli Kayıp		
		Boyut		Toplam	Boyut		Toplam	Boyut		Toplam
		1	2		1	2		1	2	
1	Cinsiyet	0.527	0.009	0.536	0.527	0.009	0.536	0.000	0.000	0.000
	Medeni hal	0.002	0.123	0.124	0.002	0.123	0.124	0.000	0.000	0.000
	Yaş Grubu	0.051	0.001	0.053	0.050	0.000	0.050	0.002	0.001	0.003
	Ruhsal hal	0.003	0.413	0.416	0.003	0.413	0.416	0.000	0.000	0.000
2	Adres	0.411	0.253	0.664						
	İkamet	0.002	0.028	0.030						
	Ekonomi	0.061	0.023	0.083	0.061	0.023	0.083	0.000	0.000	0.000
	Kardeş sayısı grubu	0.085	0.013	0.098	0.085	0.012	0.097	0.000	0.001	0.001
3	İntihar	0.117	0.265	0.382	0.117	0.265	0.382	0.000	0.000	0.000
	Cinsel Travma	0.045	0.000	0.045	0.045	0.000	0.045	0.000	0.000	0.000
	Fiziksel Travma	0.163	0.092	0.254	0.163	0.092	0.254	0.000	0.000	0.000
	Cinsel+Fiziksel Trav.	0.100	0.156	0.256	0.100	0.156	0.256	0.000	0.000	0.000

Çoklu uyum her bir değişken için çoklu kategorik koordinatların varyansına eşittir. Bu ölçüler aynı zamanda homojenlik analizindeki ayırma ölçüleri ile benzerdir. Bu değerlere bakarak hangi değişkenin ayırmada daha iyi olduğu belirlenebilir. Buna göre Tablo 6 incelendiğinde, adres değişkenine ait çoklu uyum değerleri toplamı 0.664 iken, kardeş sayısı değişkeninin çoklu uyum toplamı 0.098 olup, bu bilgi adres değişkeninin kardeş sayısı değişkeninden daha güçlü bir ayırım yaptığını göstermektedir. Bu değişkenler içerisinde en iyi ayırımı adres değişkeni yaparken, bunu sırası ile cinsiyet ve ruhsal hal değişkenlerinin izlediği gözlenmektedir. İkametgah ve cinsel travma değişkeni ise oldukça düşük ayırım gücüne sahiptir.

Tekli uyum her bir değişken için ağırlıkların karesi ile ilişkilidir ve tekli kategorik koordinatlarının karesine eşittir. Dolayısıyla ağırlıklar tekli kategorik koordinatlarının standart sapmasıdır. Buna göre tekli uyumun boyutlardaki durumu incelenerek, değişkenlerin temelde hangi boyutta ayrıldığı belirlenebilir. Birinci settaki değişkenlerden cinsiyet 1. boyutta ayrılırken, medeni hal 2. boyutta daha iyi ayırım yapabilmektedir. Aynı şekilde fiziksel travma 1. boyutta ayrılırken, ruhsal hal 2. boyutta daha iyi ayırım yapabilmektedir.

Tablo 7. Değişkenlere ait kategori koordinatları

SETLER	Değişkenler	Kategoriler	Kategori Koordinatları	
			Boyut 1	Boyut 2
SET I	Cinsiyet	Kız	-1.025	-0.133
		Erkek	0.514	0.067
	Yaş grubu	17-19	-0.608	-0.044
		20-25	0.090	-0.002
		26-30	-0.004	0.104
		31-35	-0.405	0.339
	Medeni Hal	Bekar	-0.007	0.065
		Evli	0.219	-1.900
	Ruhsal Hal	R.Var	-0.188	2.073
		R.Yok	0.018	-0.199
SET 2	Adres	Aile Yanı	0.171	-0.793
		Yurt	-0.920	0.190
		Pansiyon	0.995	0.063
		Evde Tek	0.302	1.425
		Evde Ark	0.533	0.231
	ikamet	B. Şehir	0.080	0.265
		Şehir	-0.021	-0.091
		Kasaba	-0.052	-0.224
		Koy	-0.035	0.126
	Ekonomi	Düşük	0.224	-0.104
		Orta	0.127	-0.081
		Yüksek	-0.440	0.269
	Kardeş Sayısı Grubu	0-5	-0.220	-0.071
		6-10	0.247	0.059
		11-15	0.850	0.428
SET 3	İntihar	İ. Var	-1.272	1.912
		İ.Yok	0.092	-0.138
	Cinsel Travma	C.Yok	-0.067	-0.001
		C.Var	0.668	0.015
	Fiziksel Travma	F.Yok	-0.232	-0.174
		F.Var	0.702	0.527
	Cinsel+Fiziksel Travma	C-F.Yok	-0.091	-0.114
		C-F.Var	1.096	1.370

Tablo 8. Boyutlar için özdeğerler ve kayıp değerleri

		Boyut		Toplam
		1	2	
Kayıp	Set 1	0.319	0.540	0.859
	Set 2	0.447	0.405	0.852
	Set 3	0.692	0.773	1.465
	Ortalama	0.486	0.573	1.059
Özdeğerler		0.514	0.427	
Uyum				0.941
Açıklanan uyum %		54.62	45.37	

Birinci ve ikinci boyuttaki ortalama kayıpların toplamı 1.059 olup, 2. boyuttaki kaybın (0.573) birinci boyuttaki kayıptan (0.486) daha yüksek olduğu görülmektedir. Toplam uyum olan 0.941 değerine göre 1. boyutla gerçek uyumun %54.62'si açıklanabilmektedir (0.514/0.941). Benzer şekilde 2. boyut ise gerçek uyumun %45.37'sini açıklayabilmektedir (0.427/0.941). Maksimum uyum değeri, boyut sayısına eşit olup, bunun gerçekleşmesi halinde tüm veri setleri ile boyutlar arasındaki mükemmel uyum var demektir.

Tablo 9. Değişkenlerin boyutlardaki ağırlıkları

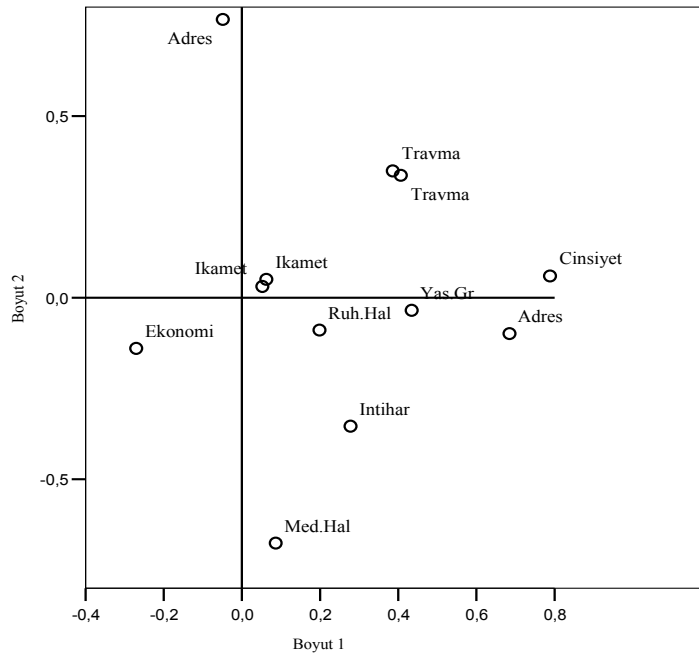
Set		Boyut	
		1	2
1	Cinsiyet	0.725	0.062
	Medeni hal	0.070	-0.675
	Yaş Grubu	0.236	-0.006
2	Ekonomi	-0.284	-0.016
3	Ruhsal hal	0.195	0.040
	İntihar	0.284	-0.333

Tablo 9 de yer alan değişkenlerin ağırlıkları incelendiğinde cinsiyet değişkeninin birinci boyuttaki ağırlığının 0.725 ikinci boyuttaki ağırlığı ise 0.062 olduğu görülmektedir. Bu da cinsiyet değişkeninin birinci boyut için güçlü bir ayırım yaptığını göstermektedir. Aynı şekilde ikinci boyuttaki -0.675 değeriyle medeni hal değişkeninin birinci boyuttaki 0.070 lik değerine göre, ikinci boyutta daha iyi bir ayırma sahip olduğu görülmektedir.

Tablo 10. Değişkenlere ait bileşen yükleri

Set				Boyut	
				1	2
1	Cinsiyet			0.789	0.060
	Medeni hal			0.087	-0.676
	Yaş Grubu			0.435	-0.035
2	Adres	Boyut	1	0.685	-0.099
			2	-0.049	0.766
	İkamet	Boyut	1	0.052	0.031
			2	0.063	0.051
	Ekonomi			-0.270	-0.139
	Ruhsal hal			0.199	-0.089
3	İntihar			0.278	-0.354
	Travma	Boyut	1	0.407	0.337
			2	0.386	0.349

Tablo 10 incelendiğinde, birinci boyut ile en yüksek korelasyona sahip olan değişken 0.789 değeri ile cinsiyet değişkenidir. Cinsiyet değişkenini 0.685 değeriyle adres değişkeni, sonrasında ise 0.435 değeri ile yaş grubu değişkeni izlemektedir. Benzer şekilde, ikinci boyutta ise en yüksek yük değeri 0.766 ile adres (2) değişkenine aittir. Bunu sırasıyla -0.676 değeri ile medeni hal değişkeni ve -0.354 değeri ile intihar değişkenlerinin izlediği görülmektedir. Değişkenlerin yük değerlerinin verilmiş olduğu Tablo 10, iki boyutlu uzayda grafiksel olarak Şekil 4’ deki gibi gösterilebilir.

**Şekil 4.** Değişkenlere ait bileşen yükleri

Bileşen yüklerine ait Şekil 4 incelendiğinde, 1. boyut; daha çok cinsiyet, adres, yaş grubu ve travma değişkenleri tarafından belirlenirken, 2. boyut ise daha çok adres, intihar ve medeni hal değişkenleri tarafından belirlenmektedir.

Tablo 11. Boyutlara ait uyum değerleri

Set		Çoklu Uyum			Tekli Uyum			Tekli Kayıp		
		Boyut		Toplam	Boyut		Toplam	Boyut		Toplam
		1	2		1	2		1	2	
1	Cinsiyet	0.526	0.004	0.530	0.526	0.004	0.530	0.000	0.000	0.000
	Medeni hal	0.005	0.456	0.461	0.005	0.456	0.461	0.000	0.000	0.000
	Yaş Grubu	0.058	0.001	0.059	0.056	0.000	0.056	0.002	0.001	0.003
2	Adres	0.479	0.590	1.069						
	İkamet	0.001	0.008	0.009						
	Ekonomi	0.081	0.001	0.082	0.081	0.000	0.081	0.000	0.000	0.001
3	Ruhsal hal	0.038	0.002	0.040	0.038	0.002	0.040	0.000	0.000	0.000
	İntihar	0.081	0.111	0.191	0.081	0.111	0.191	0.000	0.000	0.000
	Travma	0.219	0.104	0.323						

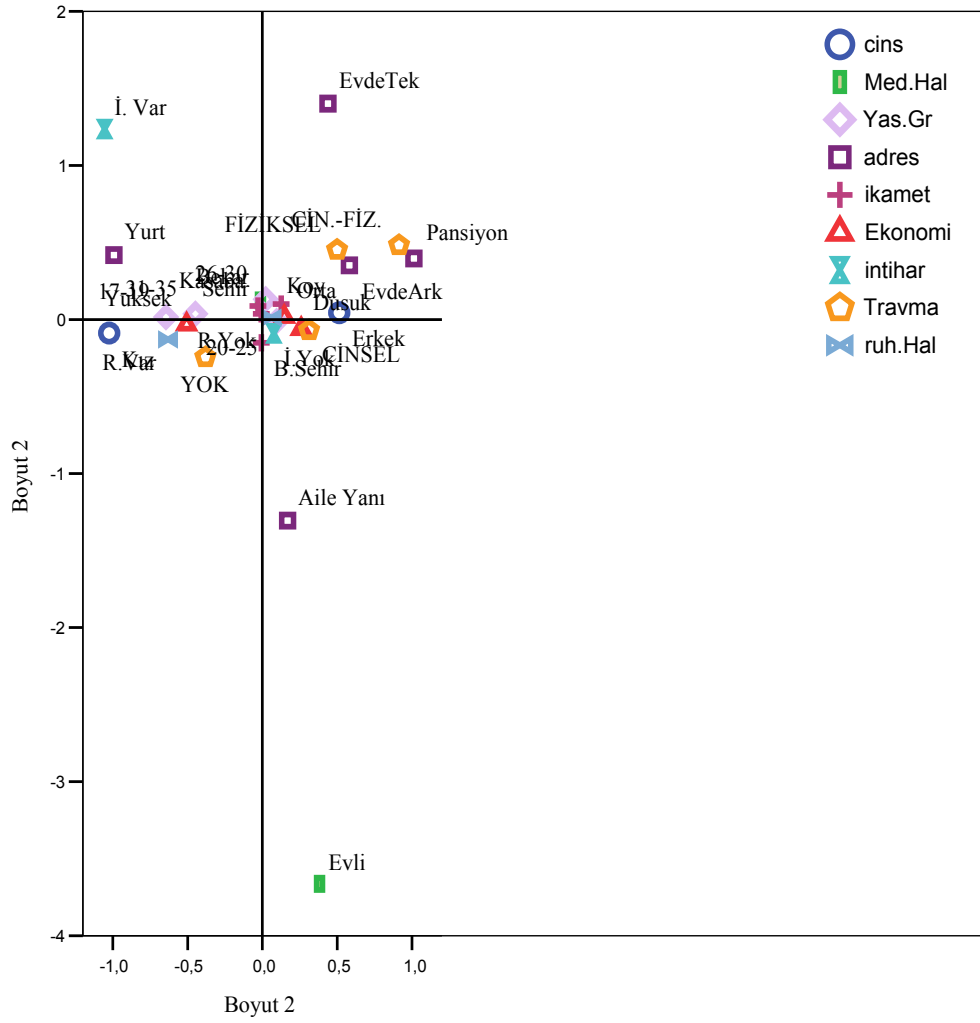
Birinci uygulamada olduğu gibi bu uygulamada da cinsiyete ait çoklu ve tekli uyum değerleri birbirine eşit çıkmıştır. Dolayısıyla, bu demektir ki çoklu koordinatlar ağırlıklarla (weights) verilen doğrular üzerindedir.

Tablo 11 incelendiğinde; adres değişkenine ait çoklu uyum değerleri toplamı 1.069 iken, ekonomi değişkeninin çoklu uyum toplamı 0.082 olup bu bilgi adres değişkeninin ekonomi değişkeninden daha güçlü bir ayırım yaptığını göstermektedir. Bu değişkenler içerisinde en iyi ayırımı adres değişkeni yaparken, bunu sırası ile cinsiyet ve medeni hal değişkenlerinin izlediği görülmektedir. İkametgâh ve ruhsal hal değişkeni ise oldukça düşük ayırım gücüne sahiptir.

Tekli uyum her bir değişken için ağırlıkların karesi ile ilişkilidir ve tekli kategorik koordinatlarının karesine eşittir. Dolayısıyla ağırlıklar tekli kategorik koordinatlarının standart sapmasıdır. Buna göre tekli uyumun boyutlardaki durumu incelenerek, değişkenlerin temelde hangi boyutta ayrıldığı belirlenebilir. Birinci setteki değişkenlerden cinsiyet 1. boyutta ayrılırken, medeni hal 2. boyutta daha iyi ayırım yapabilmektedir.

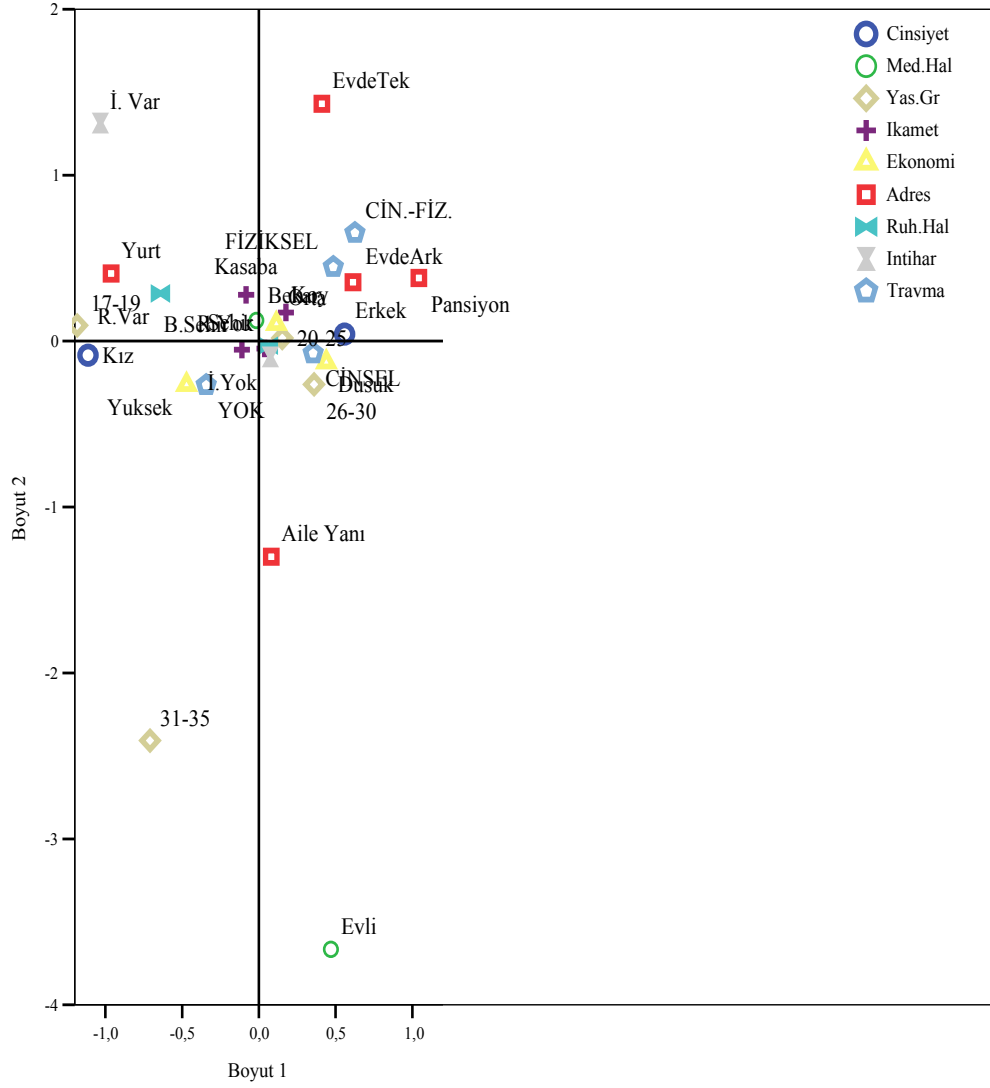
Tablo 12. Değişkenlere ait kategori koordinatları

SETLER	Değişkenler	Kategoriler	Kategori Koordinatları	
			Boyut 1	Boyut 2
SET I	Cinsiyet	Kız	-1.025	-0.088
		Erkek	0.514	0.044
	Yaş grubu	17-19	-0.644	0.016
		20-25	0.094	-0.011
		26-30	0.022	0.131
		31-35	-0.448	0.037
	Medeni Hal	Bekar	-0.013	0.124
		Evli	0.382	-3.663
SET 2	Adres	Aile Yanı	0.168	-1.305
		Yurt	-0.992	0.418
		Pansiyon	1.013	0.397
		Evde Tek	0.438	1.401
		Evde Ark	0.581	0.351
	ikamet	B. Şehir	-0.007	-0.151
		Şehir	-0.008	0.037
		Kasaba	-0.030	0.088
		Koy	0.125	0.100
	Ekonomi	Düşük	0.260	-0.061
		Orta	0.146	0.018
		Yüksek	-0.507	-0.032
SET 3	İntihar	İ. Var	-1.055	1.237
		İ.Yok	0.076	-0.090
	Travma	Cinsel Tr.	0.312	-0.070
		Fiziksel Tr.	0.499	0.452
		Cin.+Fiz. Tr.	0.913	0.482
		Travma yok	-0.381	-0.245
	Ruhsal Hal	R.Var	-0.631	-0.129
		R.Yok	0.061	0.012



Şekil 5. Değişkenlere ait çoklu kategori koordinatları

Şekil 5 incelendiğinde; köyde ikamet eden öğrencilerin, evde tek başına, arkadaşlarıyla ya da pansiyonda kaldıkları ve ekonomik düzeylerinin orta olduğu görülmektedir ve bu gruptaki öğrencilerin fiziksel ya da cinsel+fiziksel travmaya maruz kalma eğiliminde oldukları görülmektedir. 17-19 yaş ve 26-30 yaş arası, bekar, kız öğrencilerin yurttaki kaldıkları ve hayatlarının herhangi bir döneminde intiharı düşündükleri görülmüştür. Benzer şekilde, şehirde yaşayan, ekonomik düzeyi yüksek olan öğrencilerin herhangi bir travma yaşamamış olmalarına rağmen ruhsal sıkıntı yaşadıkları görülmektedir. Büyük şehirde yaşayan 20-25 yaş arası, aile yanında kalan öğrencilerin; fiziksel travma yaşama eğilimleri olmasına rağmen, hayatlarının hiçbir döneminde intiharı düşünmedikleri görülmüştür.



Şekil 6. Değişkenlerin kategorilerine ait sentroid koordinatları

Şekil 6’ da değişkenlere ait kategorilerin, değişken sentroidlerine göre dağılımı görülmektedir. Değişken kategorilerinin, sentroidlere göre dağılımı, Şekil 5’ teki kategori koordinatlarına göre dağılıma benzerdir.

4.3. Üçüncü Uygulama

Bu uygulamada, değişken seti sayısı dört olarak alınmıştır. Birinci değişken setinde; cinsiyet, yaş, medeni hal ve adres değişkenleri, ikinci değişken setinde; ikametgah, kardeş sayısı grubu ve ekonomi değişkenleri, üçüncü değişken setinde; ruhsal hal ve intihar değişkenleri, dördüncü değişken setinde ise cinsel travma, fiziksel travma ve cinsel+fiziksel travma değişkenleri yer almıştır.

Tablo 13. Boyutlar için özdeğerler ve kayıp değerleri

		Boyut		Toplam
		1	2	
Kayıp	Set 1	0.390	0.791	1.181
	Set 2	0.537	0.867	1.404
	Set 3	1.000	0.405	1.404
	Set 4	0.627	0.635	1.262
	Ortalama	0.639	0.674	1.313
Özdeğer		0.361	0.326	
Uyum				0.687
Açıklanan uyum %		52.54	47.45	

Birinci ve ikinci boyuttaki ortalama kayıpların toplamı 1.313 olarak hesaplanmıştır (Tablo 13). Birinci boyuttaki kayıp (0.639) ikinci boyuttaki kayıptan (0.674) daha düşük bulunmuştur. Özdeğerlerin her boyut için 1-ortalama kayıp değerleri de; birinci boyut için $1-0.639= 0.361$, ikinci boyut için ise $1-0.674=0.326$ dır. Birinci boyutta gerçek uyumun %52.54 ($0.361/0.687$)' ü açıklanmaktadır. İkinci boyutta ise gerçek uyumun %47.45($0.326/0.687$)' i açıklanmaktadır.

Bu uygulamada da diğer iki uygulamada olduğu gibi 2' den fazla set kullanıldığı için setler arasındaki kanonik korelasyon $r_d = ((K \times E_d)-1)/(K-1)$ eşitliği ile hesaplanmıştır. Buna göre 1. boyut için kanonik korelasyon $((4 \times 0.361)-1)/(3)= 0.148$, ikinci boyut için ise $((4 \times 0.326)-1)/(3)= 0.101$ dir.

Tablo 14. Değişkenlerin boyutlardaki ağırlıkları

Setler		1. Boyut	2. Boyut
1	Cinsiyet	0.610	0.273
	Yaş Grubu	-0.017	-0.137
	Medeni Hal	-0.028	-0.196
2	Ekonomi	-0.420	-0.198
	Kardeş Sayısı Grubu	0.485	-0.011
3	Ruhsal Hal	-0.003	0.437
	İntihar	-0.016	0.536
4	Cinsel Travma	0.024	0.063
	Fiziksel Travma	0.603	0.009
	Cinsel+Fiziksel Travma	0.258	-0.594

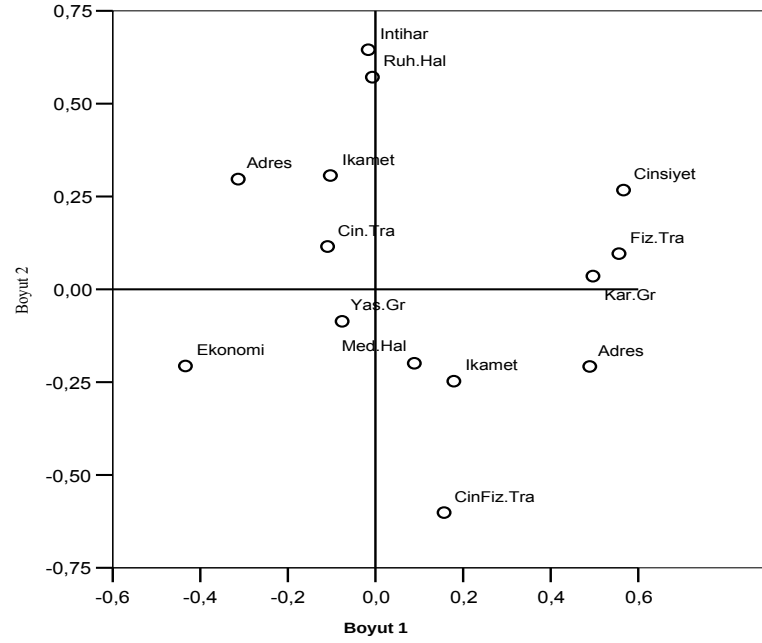
Tablo 14’ de yer alan değişkenlerin boyutlardaki ağırlıkları incelendiğinde; cinsiyet değişkeninin birinci boyuttaki ağırlığı 0.610 ikinci boyuttaki ağırlığı ise 0.273 olduğundan, cinsiyet değişkeninin birinci boyut için ikinci boyuttan daha güçlü bir ayırım yaptığı söylenebilir. Aynı şekilde ikinci boyuttaki -0.594 değeriyle; cinsel+fiziksel travma değişkeninin birinci boyuttaki 0.258 lik değerine göre, ikinci boyutta daha iyi bir ayırma sahip olduğu görülmektedir.

Tablo 15. Değişkenlere ait bileşen yükleri

Setler				Boyut 1	Boyut 2
1	Cinsiyet			0.566	0.267
	Yaş grubu			-0.076	-0.086
	Medeni Hal			0.089	-0.199
	Adres	Boyut	1	0.489	-0.208
			2	-0.314	0.297
2	İkamet	Boyut	1	0.179	-0.247
			2	-0.103	0.307
	Ekonomi			-0.434	-0.206
	Kardeş Sayısı Grubu			0.496	0.036
3	Ruhsal Hal			-0.007	0.571
	İntihar			-0.016	0.645
4	Cinsel Travma			-0.110	0.115
	Fiziksel Travma			0.556	0.096
	Cinsel+Fiziksel Travma			0.157	-0.601

Tablo 15’te değişkenlere ait nesne skorları ile dönüştürülmüş değişkenler arasındaki korelasyonları gösteren bileşen yükleri ve Şekil 7’de de Tablo 15’in iki boyutlu uzaydaki grafiksel gösterimi verilmiştir. Yükler incelendiğinde, birinci boyut ile en yüksek korelasyona sahip olan değişkenin cinsiyet değişkeni olduğu, bunu sırasıyla fiziksel travma, kardeş sayısı grubu, adres ve ekonomi değişkenlerinin izlediği görülmektedir. Benzer şekilde, ikinci boyutta ise en yüksek korelasyona intihar değişkeni sahiptir. Sonrasında ise sırasıyla ikinci boyutta en yüksek korelasyona sahip olan değişkenler, cinsel+fiziksel travma ve ruhsal hal değişkenleridir.

Değişkenlerin yük değerlerinin verilmiş olduğu Şekil 7’ de yer alan değişkenlerin bileşen yükleri incelendiğinde birinci boyut, daha çok cinsiyet, fiziksel travma, kardeş sayısı grubu, adres ve ekonomi değişkenleri tarafından belirlenirken, ikinci boyutun ise intihar, ruhsal hal ve cinsel+fiziksel travma değişkenleri tarafından belirlendiği görülmektedir.



Şekil 7. Değişkenlere ait bileşen yükleri

Tablo 16. Boyutlara ait uyum değerleri

Set		Çoklu Uyum			Tekli Uyum			Tekli Kayıp		
		Boyut		Toplam	Boyut		Toplam	Boyut		Toplam
		1	2		1	2		1	2	
1	Cinsiyet	0.372	0.075	0.446	0.372	0.075	0.446	0.000	0.000	0.000
	Yaş Grubu	0.001	0.020	0.021	0.000	0.019	0.019	0.001	0.001	0.002
	Medeni Hal	0.001	0.038	0.039	0.001	0.038	0.039	0.000	0.000	0.000
	Adres	0.295	0.083	0.378						
2	İkamet	0.049	0.091	0.140						
	Ekonomi	0.177	0.039	0.216	0.177	0.039	0.216	0.000	0.000	0.000
	Kardeş Sayısı Grubu	0.235	0.000	0.235	0.235	0.000	0.235	0.000	0.000	0.000
3	Ruhsal Hal	0.000	0.191	0.191	0.000	0.191	0.191	0.000	0.000	0.000
	İntihar	0.000	0.287	0.287	0.000	0.287	0.287	0.000	0.000	0.000
4	Cinsel Travma	0.001	0.004	0.004	0.001	0.004	0.004	0.000	0.000	0.000
	Fiziksel Travma	0.363	0.000	0.364	0.363	0.000	0.364	0.000	0.000	0.000
	Cinsel+Fiziksel Travma	0.067	0.353	0.420	0.067	0.353	0.420	0.000	0.000	0.000

Diğer iki uygulamada olduğu gibi bu uygulamada da cinsiyet değişkenine ait çoklu ve tekli uyum değerleri birbirine eşit çıkmıştır. Bu nedenle de, çoklu koordinatlar ağırlıklarla (weights) verilen doğrular üzerindedir denilebilir.

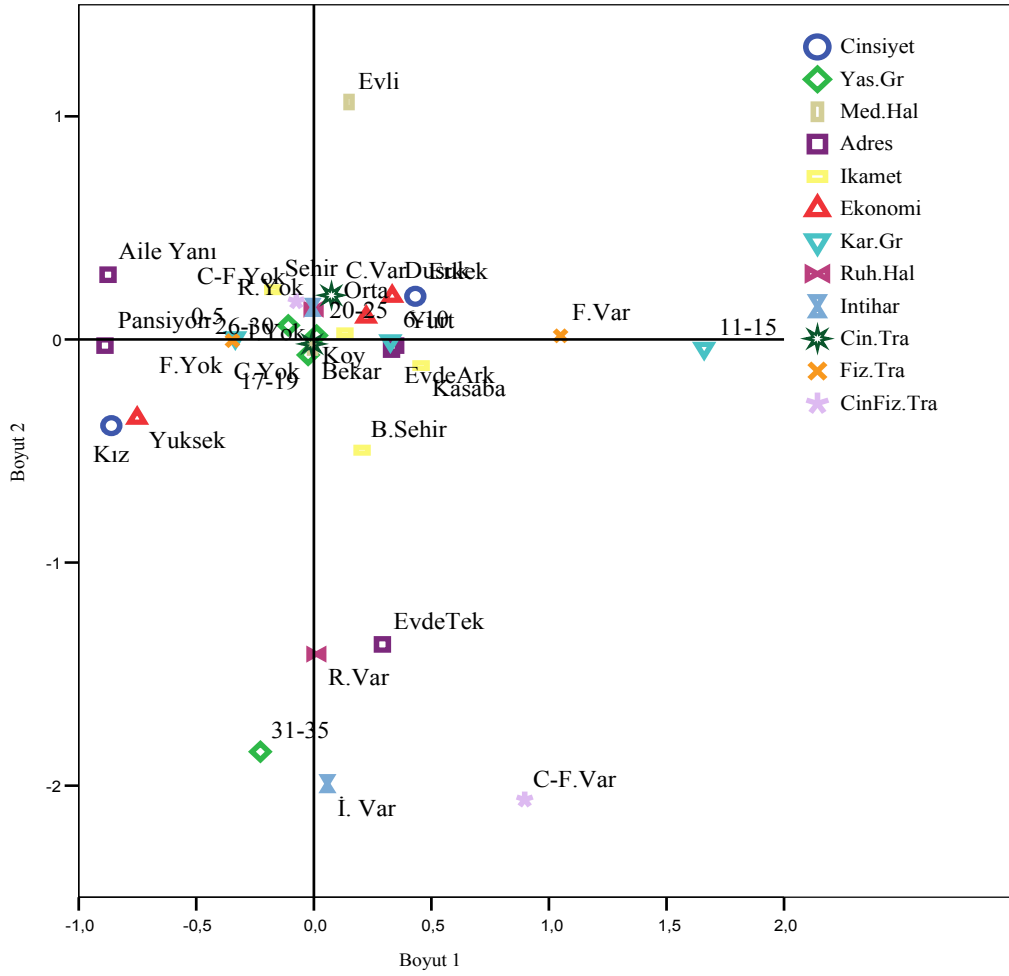
Tablo 16’ daki değerlere bakarak bu analiz sonucunda hangi değişkenin ayırmada daha iyi olduğu belirlenebilir. Buna göre Tablo 16 incelendiğinde, cinsiyet değişkenine ait çoklu uyum değerleri toplamı 0.446 iken, cinsel travma değişkeninin çoklu uyum toplamı 0.004 olup, bu bilgi cinsiyet değişkeninin cinsel travma değişkeninden daha güçlü bir ayırım yaptığını göstermektedir. Bu değişkenler içerisinde en iyi ayırımı cinsiyet değişkeni yaparken, bunu sırası ile, 0.420 değeriyle cinsel+fiziksel travma değişkeni, 0.378 değeriyle adres değişkeni ve 0.364 değeriyle de fiziksel travma değişkeni takip etmektedir. Medeni hal (0.039), yaş grubu (0.021) ve cinsel travma (0.004) değişkenlerinin ise oldukça düşük ayırım gücüne sahip oldukları görülmektedir.

Daha önceki iki uygulamada da belirtildiği gibi, tekli uyumun boyutlardaki durumu incelenerek, değişkenlerin temelde hangi boyutta ayrıldığı belirlenebilir. Bu uygulamada da birinci setteki değişkenlerden; 0.372 değeriyle cinsiyet, 0.177 değeriyle ekonomi, 0.235 değeriyle, kardeş sayısı grubu ve 0.363 değeriyle fiziksel travma değişkenleri 1. boyutta ayrılırken, 0.019 değeriyle yaş grubu, 0.191 değeriyle ruhsal hal, 0.038 değeriyle medeni hal, 0.287 değeriyle intihar, 0.004 değeriyle cinsel travma ve 0.353 değeriyle cinsel+fiziksel travma değişkenleri 2. boyutta daha iyi ayırım yapabilmektedir.

Tablo 17’de değişkenlere ait kategori koordinatları verilmiştir. Tablo 17 incelendiğinde; cinsiyet değişkeninin “Kız” kategorisinin birinci boyutta daha etkili olduğu görülürken, medeni hal değişkeninin her iki kategorisinin de ikinci boyutta daha etkili olduğu görülmektedir. Bu koordinatların grafiksel olarak gösterimi ise Şekil 8’de verilmiştir

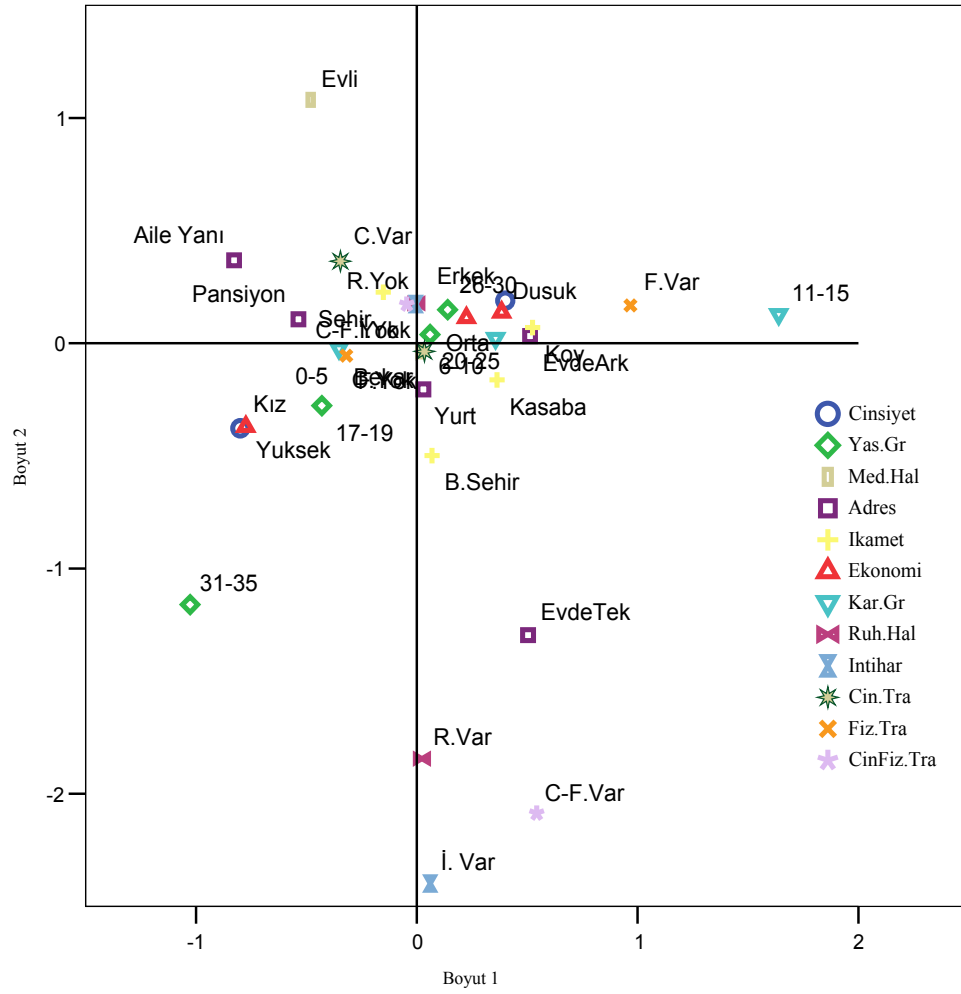
Tablo 17. Değişkenlere ait kategori koordinatları

SETLER	Değişkenler	Kategoriler	Kategori Koordinatları	
			Boyut 1	Boyut 2
SET I	Cinsiyet	Kız	-0.861	-0.386
		Erkek	0.432	0.194
	Yaş grubu	17-19	-0.024	-0.070
		20-25	0.012	0.018
		26-30	-0.108	0.063
		31-35	-0.227	-1.848
	Medeni Hal	Bekar	-0.005	-0.036
		Evli	0.150	1.064
	Adres	Aile Yanı	-0.875	0.290
		Yurt	0.348	-0.027
		Pansiyon	-0.888	-0.027
		EvdeTek	0.292	-1.367
EvdeArk		0.331	-0.045	
Set 2	İkamet	B.Şehir	0.205	-0.496
		Şehir	-0.175	0.223
		Kasaba	0.456	-0.117
		Koy	0.132	0.030
	Ekonomi	Düşük	0.334	0.194
		Orta	0.223	0.100
		Yüksek	-0.752	-0.353
	Kardeş Sayısı Grubu	0-5	-0.334	0.007
6-10		0.326	-0.006	
11-15		1.660	-0.040	
Set 3	Ruhsal Hal	R.Var	0.011	-1.410
		R.Yok	-0.001	0.135
	İntihar	İ. Var	0.058	-1.991
		İ.Yok	-0.004	0.144
Set 4	Cinsel Travma	C.Yok	-0.008	-0.020
		C.Var	0.075	0.198
	Fiziksel Travma	F.Yok	-0.346	-0.005
		F.Var	1.049	0.016
	Cinsel+Fiziksel Travma	C+F.Yok	-0.074	0.171
		C+F.Var	0.897	-2.062



Şekil 8. Değişkenlere ait çoklu kategori koordinatları

Şekil 8, incelendiğinde; 6-10 adet kardeşe sahip, erkek ve evli olan, köyde yaşayan, gelir düzeyi düşük veya orta olan öğrencilerin yurttan kalmayı tercih ettikleri ve yine bu gruptaki öğrencilerin cinsel veya fiziksel travmaya maruz kaldıkları ancak intiharı düşünmedikleri görülmektedir. Bununla birlikte kasaba ya da büyük şehirde yaşayan, evde tek başına ya da arkadaşlarıyla kalan, 11-15 adet kardeşe sahip olan, 20-25 yaş arası bireylerin, daha çok cinsel+ fiziksel travmaya maruz kalma eğiliminde oldukları ve ruhsal bunalım yaşayıp intiharı düşündükleri görülmektedir. Benzer şekilde şehirde yaşayan, 26- 30 yaş arası olup, aile yanında kalan, bekar, bireylerin cinsel+ fiziksel travma yaşamadıkları ve ruhsal bunalım geçirmediği ortaya çıkmıştır. Şekil 8'den, 0-5 adet kardeşe sahip, 17-19 ve 31-35 yaş arası, gelir düzeyi yüksek olan kız öğrencilerin yurttan kalmayı tercih ettikleri ve yine bu öğrencilerin cinsel veya fiziksel travmaya maruz kalmadıkları da söylenebilir.



Şekil 9. Değişkenlerin kategorilerine ait sentroid koordinatları

Şekil 9’da değişkenlere ait kategorilerin, değişken sentroidlerine göre dağılımı görülmektedir. Şekil 9’daki değişkenlerin kategorilerinin sentroidlere göre dağılımı, Şekil 8’de yer alan çoklu kategori koordinatlarının dağılımına benzerlik göstermektedir.

4.4. Dördüncü Uygulama

Bu uygulamada değişken seti sayısı iki olarak belirlenmiştir. Birinci değişken setine; cinsiyet, yaş, medeni hal, adres, ikametgah, kardeş sayısı grubu ve ekonomi değişkenleri alınırken, ikinci değişken setine de ruhsal hal, intihar, cinsel travma, fiziksel travma ve cinsel+fiziksel travma değişkenleri alınmıştır.

Tablo 18. Boyutlar için özdeğerler ve kayıp değerleri

		Boyut		Toplam
		1	2	
Kayıp	Set 1	0.351	0.397	0.748
	Set 2	0.351	0.397	0.749
	Ortalama	0.351	0.397	0.748
Özdeğer		0.649	0.603	
Uyum				1.252
Açıklanan uyum %		51.83	48.16	

Tablo 18 incelendiğinde, birinci ve ikinci boyuttaki ortalama kayıpların toplamı 0.748 olarak hesaplanmıştır. Birinci boyuttaki kayıp (0.351) ikinci boyuttaki kayıptan (0.397) daha küçük bulunmuştur. Özdeğerlerin her boyut için 1-ortalama kayıp değerleri de, birinci boyut için $1-0.351=0.649$, ikinci boyut için ise $1-0.397=0.603$ olarak hesaplanmıştır.

Birinci boyutta gerçek uyumun %51.83 ü açıklanabilmekte iken ($0.649/1.252$), ikinci boyutta ise gerçek uyumun ancak %48.16 sı açıklanmaktadır ($0.603/1.252$).

Bundan önceki üç uygulamada da 2 den fazla değişken seti kullanıldığı için setler arasındaki kanonik korelasyon $r_d = ((K \times E_d) - 1) / (K - 1)$ eşitliği ile hesaplanarak bulunmuştu. Ancak bu uygulamada ise 2 değişken seti kullanıldığı için setler arasındaki kanonik korelasyon $r = 2 \times E_d - 1$ eşitliği ile hesaplanabilir. Buna göre 1. boyut için kanonik korelasyon $2 \times 0.649 - 1 = 0.298$ olarak hesaplanırken, İkinci boyut için ise $2 \times 0.603 - 1 = 0.206$ olarak bulunur.

Tablo 19. Değişkenlerin boyutlardaki ağırlıkları

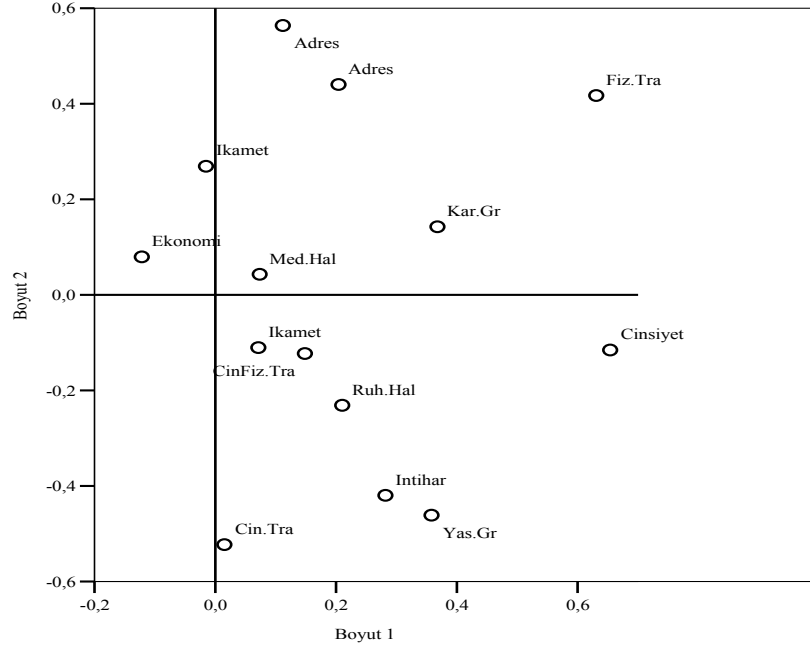
Set		Boyut	
		1	2
1	Cinsiyet	0.602	0.097
	Medeni Hal	0.126	0.185
	Ekonomi	-0.017	0.131
	Yas Grubu	0.188	-0.464
	Kardeş sayısı Grubu	0.290	0.125
2	Ruhsal Hal	0.230	-0.124
	İntihar	0.283	-0.421
	Cinsel Travma	0.168	-0.479
	Fiziksel Travma	0.733	0.291
	Cinsel+Fiziksel Travma	0.374	-0.211

Tablo 19’da yer alan değişkenlerin ağırlıkları incelendiğinde fiziksel travma değişkeninin birinci boyuttaki ağırlığı 0.733, ikinci boyuttaki ağırlığı ise 0.291 olduğu için fiziksel travmanın birinci boyuttaki etkisinin ikinci boyuta göre daha fazla olduğu söylenebilir. Aynı şekilde cinsiyet, cinsel+fiziksel travma, kardeş sayısı grubu ve ruhsal hal değişkenlerinin birinci boyuttaki etkilerinin, ikinci boyuta göre daha fazla olduğundan birinci boyut için daha iyi bir ayırım yaptıkları görülmektedir. Benzer şekilde cinsel travma değişkeninin ikinci boyuttaki ağırlığının -0.479 birinci boyuttaki ağırlığının ise 0.168 olması cinsel travma değişkeninin ikinci boyut için daha iyi bir ayırım yaptığını göstermektedir. Yaş grubu, intihar, ekonomi ve medeni hal değişkenleri de ağırlıkları bakımından ikinci boyutta daha etkili olmuşlardır.

Tablo 20. Değişkenlere ait bileşen yükleri

Set				Boyut	
				1	2
1	Cinsiyet			0.654	-0.116
	Medeni hal			0.074	0.043
	Adres	Boyut	1	0.204	0.440
			2	0.112	0.564
	İkamet	Boyut	1	0.071	-0.110
			2	-0.015	0.269
	Ekonomi			-0.121	0.079
	Yas Grubu			0.358	-0.461
Kardeş sayısı Grubu			0.368	0.142	
2	Ruhsal Hal			0.210	-0.231
	İntihar			0.282	-0.420
	Cinsel Travma			0.015	-0.523
	Fiziksel Travma			0.631	0.417
	Cinsel+Fiziksel Travma			0.148	-0.123

Tablo 20’deki bileşen yükleri incelendiğinde birinci boyut ile en yüksek korelasyona cinsiyet değişkeninin sahip olduğu, bunu sırasıyla fiziksel travma ve kardeş sayısı grubu değişkenlerinin izlediği görülmektedir. Benzer şekilde, ikinci boyutta ise en yüksek korelasyona adres, cinsel travma, yaş grubu ve intihar değişkenlerinin sahip olduğu görülmektedir. Şekil 10’ da ise Tablo 20’deki bileşen yüklerinin, iki boyutlu uzayda grafiksel olarak gösterimi verilmiştir. Şekil 10’dan değişkenlerin orijine olan uzaklıkları dikkate alınarak, bileşen yüklerinin boyutlar üzerindeki etkisi görülebilir.



Şekil 10. Değişkenlere ait bileşen yükleri

Şekil 10’ da yer alan değişkenlerin bileşen yükleri incelendiğinde; birinci boyutta, cinsiyet ve fiziksel travma değişkenlerinin etkili olduğu görülürken, ikinci boyutta ise cinsel travma, yaş grubu ve adres değişkenlerinin etkili olduğu görülmüştür.

Tablo 21. Boyutlara ait uyum değerleri

Set		Çoklu Uyum			Tekli Uyum			Tekli Kayıp		
		Boyut		Toplam	Boyut		Toplam	Boyut		Toplam
		1	2		1	2		1	2	
1	Cinsiyet	0.363	0.009	0.372	0.363	0.009	0.372	0.000	0.000	0.000
	Medeni Hal	0.016	0.034	0.050	0.016	0.034	0.050	0.000	0.000	0.000
	Adres	0.084	0.306	0.389						
	İkamet	0.023	0.038	0.061						
	Ekonomi	0.000	0.017	0.017	0.000	0.017	0.017	0.000	0.000	0.000
	Yaş Grubu	0.037	0.216	0.253	0.035	0.215	0.250	0.002	0.001	0.003
	Kardeş Grubu	0.084	0.016	0.100	0.084	0.016	0.099	0.000	0.000	0.000
2	Ruhsal Hal	0.053	0.015	0.068	0.053	0.015	0.068	0.000	0.000	0.000
	İntihar	0.080	0.177	0.257	0.080	0.177	0.257	0.000	0.000	0.000
	Cinsel Travma	0.028	0.229	0.257	0.028	0.229	0.257	0.000	0.000	0.000
	Fiziksel Travma	0.537	0.085	0.622	0.537	0.085	0.622	0.000	0.000	0.000
	Cinsel+ Fiziksel Travma	0.140	0.045	0.184	0.140	0.045	0.184	0.000	0.000	0.000

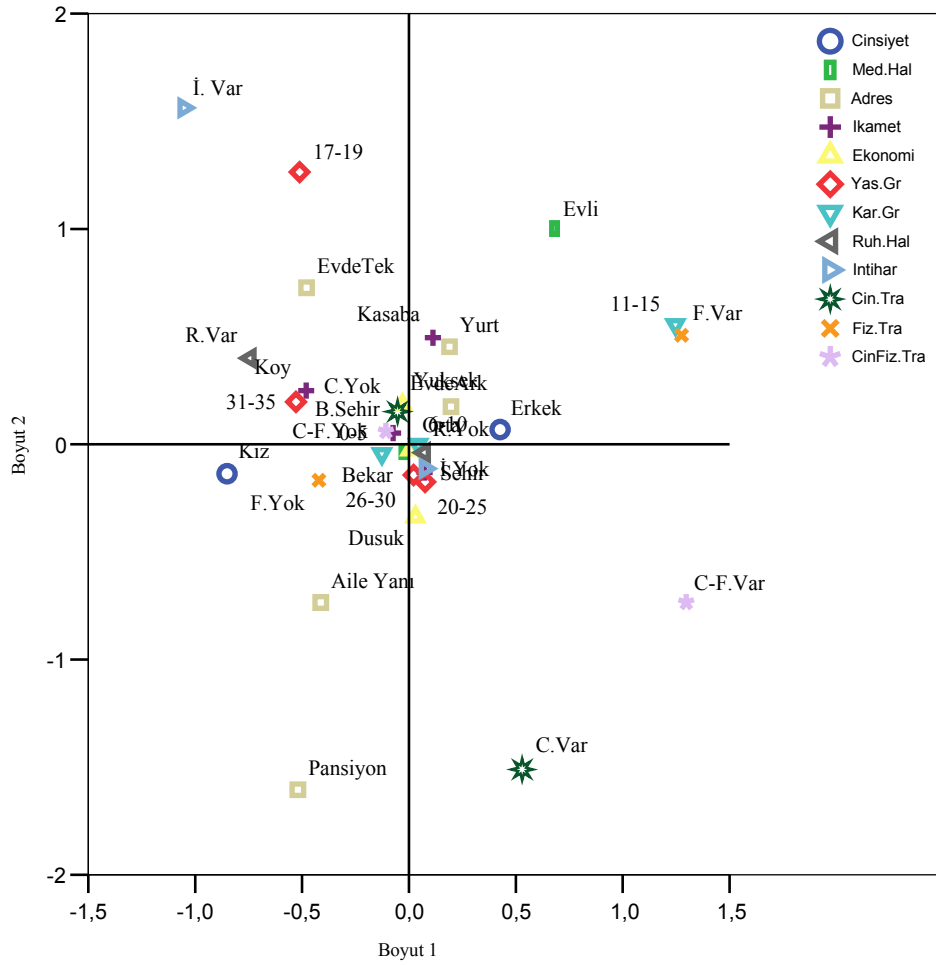
Tablo 21 incelendiğinde, fiziksel travma değişkenine ait çoklu uyum değerleri toplamının 0.622, ekonomi değişkeninin çoklu uyum toplamının da 0.017 olduğu dikkat çekmektedir. Bu da fiziksel travma değişkeninin ekonomi değişkeninden daha güçlü bir ayırım yaptığını göstermektedir. Bu uygulamada kullanılan değişkenler içerisinde en iyi ayırımın fiziksel travma değişkeni ile belirlendiği ve bu değişkeni sırası ile adres ve cinsiyet değişkenlerinin takip ettiği görülmektedir. Bunun yanı sıra medeni hal, ikametgah ve ekonomi değişkenlerinin ise oldukça düşük ayırım gücüne sahip olduğu dikkat çekmektedir.

Yine bu analiz sonucunda da, tekli uyumun boyutlardaki durumu incelenerek, değişkenlerin temelde hangi boyutta ayrıldığı belirlenebilir. Bu uygulamada da kullanılan değişkenlerden cinsiyet, kardeş sayısı grubu, ruhsal hal, fiziksel travma ve cinsel+fiziksel travma 1. boyutta ayrılırken, yaş grubu, ekonomi, medeni hal, intihar ve cinsel travma değişkenleri 2. boyutta daha iyi ayırım yapmıştır.

Tablo 22’de değişkenlere ait kategori koordinatları verilmiştir. Tablo 22 incelendiğinde; 1. setteki cinsiyet değişkeninin kız ve erkek kategorileri ile yaş değişkeninin 31-35 kategorisi birinci boyutta etkili olurken, yaş değişkeninin 17-19, 20-25 ve 26-30 kategorileri 2. boyutta daha etkili olmuştur. Aynı şekilde medeni hal değişkeninin her iki kategorisi de 2. boyutta daha etkilidir. Adres değişkeninin “evde arkadaşlarıyla kategorisi” 1. boyutta, “aile yanı”, “yurt”, “pansiyon” ve “evde tek” kategorileri ise 2. boyutta daha etkilidir. İkamet değişkeninin büyük şehir ve köy kategorileri 1. boyutta, şehir ve kasaba kategorileri ise 2. boyutta etkilidir. Ekonomi değişkeninin bütün kategorileri 2. boyutta etkili olurken, kardeş sayısı grubu değişkeninin bütün kategorileri 1. boyutta daha etkili olmuştur. Benzer şekilde ikinci setteki değişkenlerden ruhsal hal, fiziksel travma ve cinsel+fiziksel travma değişkenlerinin her iki kategorileri de 1. boyutta etkili olurken, intihar ve cinsel travma değişkenlerinin kategorileri ise 2. boyutta etkili olmuştur. Tablo 22’de rakamsal olarak verilen kategori koordinatlarının grafiksel olarak gösterimi Şekil 11’ de verilmiştir.

Tablo 22. Değişkenlere ait kategori koordinatları

SETLER	Değişkenler	Kategoriler	Kategori Koordinatları	
			Boyut 1	Boyut 2
SET I	Cinsiyet	Kız	-0.851	-0.137
		Erkek	0.426	0.069
	Yaş grubu	17-19	-0.511	1.264
		20-25	0.076	-0.174
		26-30	0.022	-0.143
		31-35	-0.528	0.197
	Medeni Hal	Bekar	-0.023	-0.034
		Evli	0.681	1.002
	Adres	Aile Yanı	-0.413	-0.735
		Yurt	0.190	0.453
		Pansiyon	-0.520	-1.605
		Evde Tek	-0.479	0.727
		Evde Ark	0.196	0.175
	İkamet	B. Şehir	-0.074	0.052
		Şehir	0.073	-0.131
		Kasaba	0.112	0.495
		Koy	-0.481	0.249
	Ekonomi	Düşük	0.030	-0.337
		Orta	0.007	-0.026
		Yüksek	-0.031	0.183
	Kardeş Sayısı Grubu	0-5	-0.127	-0.044
		6-10	0.043	-0.001
		11-15	1.247	0.556
SET 2	Ruhsal Hal	R.Var	-0.744	0.400
		R.Yok	0.071	-0.038
	İntihar	İ. Var	-1.053	1.563
		İ.Yok	0.076	-0.113
	Cinsel Travma	C.Yok	-0.053	0.152
		C.Var	-0.530	-1.510
	Fiziksel Travma	F.Yok	-0.421	-0.167
		F.Var	1.276	0.506
	Cinsel+Fiziksel Travma	C+F.Yok	-0.108	0.061
		C+F.Var	1.297	-0.733



Şekil 11. Değişkenlere ait çoklu kategori koordinatları

Şekil 11 incelendiğinde; 6-10 veya 11-15 adet kardeşe sahip, erkek, kasabada yaşayan ve evli olan, gelir düzeyi orta olan öğrencilerin, yurttta veya evde arkadaşlarıyla kalmayı tercih ettikleri ve yine bu gruptaki öğrencilerin fiziksel travmaya maruz kaldıkları görülmektedir. Benzer şekilde, köyde ya da büyük şehirde yaşayan, evde tek başına kalan, gelir düzeyi yüksek, 17-19 ve 31-35 yaş arası bireylerin, herhangi bir travmaya maruz kalmadıkları, ancak ruhsal bunalım yaşayıp intiharı düşündükleri belirlenmiştir. Bununla birlikte aile yanında kalan, bekar, 26- 30 yaş arası kız öğrencilerin fiziksel travma yaşamadıkları ve pansiyonda kalmayı tercih ettikleri görülmüştür. Yine şekil 11'den 20-25 yaş arası, gelir düzeyi düşük olan ve şehirde yaşayan öğrencilerin, cinsel veya cinsel+fiziksel travmaya maruz kaldıkları ancak intiharı düşünmedikleri görülmektedir.

5. SONUÇ

Uygulamada, araştırmacıların inceleme konusu olan değişkenler, genelde diğer birçok değişken ile ilişkilidir. Dolayısıyla, değişkenler arasındaki ilişkiyi incelemede, çok sayıdaki değişkenin ele alınması, gerçek ilişki yapısının bozulmaması açısından önemlidir.

Bilindiği üzere, standart kanonik korelasyon analizi için setlerde yer alan değişkenlerin nümerik olması koşulu vardır. Eğer setlerde yer alan değişkenler nümerik değilse ve setler arasında doğrusal ilişkilerin yanı sıra doğrusal olmayan ilişkiler de araştırılmak isteniyorsa, bu durumda standart kanonik korelasyon analizinin etkinliği azalır. Bu nedenle, doğrusal olmayan kanonik korelasyon analizinin kullanılması daha uygun olacaktır. Doğrusal olmayan kanonik korelasyon analizi değişkenlerin nominal ve ordinal olması durumunda da kullanılabilir. Süreçte değişkenlerin kategorileri korunarak doğrusal olmayan dönüşüm (transformasyon) yapılır ve sonuçlar değişmez. Teknik optimal ölçekleme ilkesine dayanır. Yani kategorik değişkenler, her kategorik değişken için atanan rakamsal değerlere (skor değerlerine) dönüştürülür. Bu skor değerleri, değişken setleri arasındaki ilişkiyi optimum yapacak şekildedir. Çözüm algoritması, iteratif çözümü gerektirir ve genellikle dalgali en küçük kareler algoritması olarak bilinen çözüm yöntemi kullanılır. Bu algoritma yeniden ölçeklendirilmiş bağımlı ve bağımsız değişken setlerinin doğrusal kombinasyonları arasındaki farkların kareler toplamını minimum yapacak şekilde kayıp fonksiyonunu kullanır. Kategori skorları ve değişken ağırlıkları birbirini takip edecek şekilde her veri seti için düzeltilir ve algoritma tekrarlamalı olarak devam eder. Diğer değişken kümeleri ile ilişkili olarak, ilgili değişken kümesindeki bazı sınırlamaları gerektirir. Bu bağlamda Gram-Schmit yöntemi kullanılır. Algoritma yakınsama algoritmasıdır. Çünkü hedefteki kayıp fonksiyonun değeri herhangi bir yeniden ölçekleme ile asla yükselmez.

Doğrusal Olmayan Kanonik Korelasyon Analizi, optimal ölçekleme ile yapılan kategorik kanonik korelasyon analizi ile eşdeğer olup, kısaca OVERALS olarak adlandırılır. OVERALS işlem süreci DOKKA'nın işlem süreci olarak adlandırılır. Setlerdeki değişkenler nominal, ordinal veya aralıklı ölçekte olabilir. İki'den fazla

değişken setleri arasındaki ilişki yapısı incelenebilir. Bu değişken setlerinden birisi bağımlı değişken seti, diğerleri ise bağımsız değişken setleri olarak düşünülebilir. Standart CCA’de değişken setleri arasındaki kanonik korelasyon maksimize edilmeye çalışılırken, DOKKA’ da nesne skorları ile tanımlanmış olan değişkenlerin karşılaştırılması yapılır. Nominal değişkenler için kategorilerin sırası aynı kalmaz fakat her kategori için uyum iyiliğini maksimize edecek şekilde yeniden değerler elde edilir. Sıralı değişkenler için kategorilerin sırası aynı tutulur ve uyumu maksimize edecek şekilde yeni değerler elde edilir. Sürekli değişkenler için de değerler arasındaki uzaklıklar ve sırası aynı kalacak şekilde yeniden ölçekleme yapılır.

DOKKA’nın Kanonik korelasyon ve alternatif olarak kullanılabilecek diğer analiz yöntemlerine göre üç önemli avantajının olduğu söylenebilir. Bunlardan birincisi, Optimal ölçekleme, dolayısıyla DOKKA, ikiden fazla değişken kümesi arasındaki ilişkinin bulunmasını da sağlar. İkincisi, kümelerde yer alan değişkenler; istenilen ölçekte (nominal, ordinal ya da nümerik) olabilir ve üçüncüsü doğrusal olmayan ilişkiler de analiz edilebilir. Optimal Ölçekleme ile Kategorik Kanonik Korelasyon Analizi veya DOKKA, değişken kümeleri arasındaki ilişkilerin istenilen boyutta grafiksel olarak gösterilmesine de olanak sağlamaktadır Bunların yanı sıra iki tane önemli dezavantajının da olduğu söylenebilir. Bunlardan birincisi, sürekli değişkenler kategorize edilerek analize dahil edilmektedir. Dolayısıyla bu durum, bir miktar bilgi kaybı olarak düşünülebilir. İkincisi ise herhangi bir istatistik önemlilik (anlamlılık) testinin yapılamamasıdır. Bu bağlamda, DOKKA doğrulayıcı analiz tekniklerinden daha çok araştırmacı analiz tekniklerinden biri olarak görülmektedir. Diğer yandan, Kanonik Korelasyondaki değişken kümeleri arasındaki korelasyonun maksimum yapılması yerine, DOKKA’ da deney üniteleri ya da nesnelere ait skor değerleri arasındaki uyum karşılaştırılır.

DOKKA hakkında genel bilgilerin sunulması, diğer analizlerle olan ilişkilerine değinilmesi ve elde edilen sonuçların yorumlanması açısından yapılan bu çalışmanın; ileride konu ile ilgilenecek olan araştırmacılara yardımcı olabileceği ümit edilmektedir.

ÖZET

Gönül, S, Doğrusal olmayan kanonik korelasyon analizi ve sağlık alanında Uygulaması. Yüzüncü Yıl Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı Yüksek Lisans Tezi, Van, 2010. Doğrusal Olmayan Kanonik Korelasyon Analizi, değişken kümesi sayısının 2 veya 2' den fazla ve değişkenlerin sürekli, sıralı veya kategorik olması durumunda; bu değişken kümeleri arasındaki doğrusal ve doğrusal olmayan ilişkiyi incelemeye kullanılan çok değişkenli istatistik analiz yöntemlerinden birisidir. Doğrusal Olmayan Kanonik Korelasyon Analizi, değişken kümelerini oluşturan değişkenlerin, sürekli ve aralarındaki ilişkinin de doğrusal olması gerekliliği varsayımını içermez. Dolayısıyla, değişken tipi ve aralarındaki ilişki bakımından kanonik korelasyon analizine göre oldukça esnektir. Bu çalışmada, doğrusal olmayan kanonik korelasyon analizi, genel olarak tanıtılmış, teorik alt yapısı açıklanmış ve konunun anlaşılmasını kolaylaştırmak amacıyla bir uygulama yapılmıştır. Uygulamada farklı değişken kombinasyonları ele alınarak elde edilen sonuçlar tablolar ve grafikler halinde sunulmuş ve yorumlanmıştır.

Anahtar kelimeler: Optimal ölçekleme, kayıp fonksiyonu, bileşen yükü, sentroid.

SUMMARY

Gönül, S, Nonlinear canonical correlation analysis and application in health science. Yuzuncu Yil University, Institute of Health Sciences, Master Thesis in Department of Biostatistics, Van, 2010.

Nonlinear Canonical Correlation Analysis is one of the multivariate analysis techniques which is used to examine linear and nonlinear relationships among two or more variable sets included categorical, ordinal or numerical variables. Linearity assumption does not need for Nonlinear Canonical Correlation Analysis. Thus, Nonlinear Canonical Correlation Analysis is more flexible than canonical correlation analysis in terms of relationships among variables and variable type. In this study, Nonlinear Canonical Correlation Analysis was introduced and explained theoretical basis and then done an application to expedite of the subject. In the application various variable combinations were considered and obtained results were presented as tables and graphics then these were interpreted.

Key words: Optimal scaling, loss function, component loading, centroid.

KAYNAKLAR

- Anonymous (2007). SPSS Inc. Polar Engineering and Consulting, USA.
- Bayram N, Ertay S (2001). Tüketim harcamaları davranış biçimi: Princals ve Overals Yaklaşımı. V. Ulusal Ekonometri ve İstatistik Sempozyumu, Adana.
- Bohacek S, Jonckheere E (1998). Chaotic Modeling in Network Spinal Analysis: Nonlinear Canonical Corelation with Alternating Conditional Expectation (ACE): A Preliminary Report. Request for reprints: Requests should be made to Stephen Bohack, department, Department of Electrical Engineering-Systems. Universty of Southern California, Los Angeles, U.S.A.
- Boysan M, Goldsmith RE, Çavuş H, Kayri M, Keskin S (2009). Relations among anxiety, Depression and Dissociative Symptoms. The Influence of Abuse Subtype. Journal of Trauma and Dissociatio, 10, 83-101.
- Cannon AL (2006). Nonlinear principal predictor analysis: Application to the lorenz system. Journal Of Climate, 19, 579-589.
- Cannon AC, Hsieh WW (2008). Robust nonlinear canonical correlation analysis: application to seasonal climate forecasting. Nonlin. Processes Geophys, 15, 221–232.
- Frie KG, Janssen C (2009). Department of Medical Sociology, Center of Psychosocial Medicine, University Medical Center Hamburg-Eppendorf, Martinistr, 52, 20246 Hamburg, Germany. k.grosse-frie@uke.uni-hamburg.de Erişim Tarihi: 23.04.2010.
- Gifi A (1990). Nonlinear Multivariate Analysis. John Wiley&Sons, New York.
- Girginer N, Kaygısız Z, Yalama A (2007). Doğrusal olmayan kanonik korelasyon analizi ile istatistiğe yönelik tutumlarda üniversite öğrencileri arasındaki bireysel farklılıkların incelenmesi, İstanbul Üniversitesi İktisat Fakültesi Ekonometri ve İstatistik Dergisi, 6 (1), 29-40.
- Golob TF (1985). Nonlinear Canonical Correlation Analysis of Weekly Trip Chaining Behaviour. Institute of Transportation Studies University of California, Irvine, CA, U.S.A.
- Gürbüz F. (1989). Değişken takımları arasındaki ilişkinin kanonik korelasyon analizi ile araştırılması Ankara Üniveristesi, Ziraat Fakültesi, Yayın No 1162. Ankara.
- Hsieh WW (2000). Nonlinear Canonical Correlation Analysis of the Tropical Pacific Climate Variability Using a Neural Network Approach. Oceanography/EOS, University of British Columbia, Vancouver, British Columbia, Canada.
- Hsieh WW (2001). Nonlinear Canonical Correlation Analysis by Neural Networks Department of Earth and Ocean Sciences. University of British Columbia,Vancouver, BC, Canada.

Keskin S (2002). Varyansların homojenliğini test etmede kullanılan bazı yöntemlerin I. tip hata ve testin gücü bakımından irdelenmesi. Ankara Üniversitesi, Fen Bilimleri Enstitüsü. Zootekni Anabilim Dalı, Doktora Tezi, Ankara.

Keskin S, Özsoy A (2004). Kanonik korelasyon analizi ve bir uygulaması. Tarım Bilimleri Dergisi, 10(1), 67-71.

Keskin S., Ankaralı H, Noyan T, Kamacı M (2007). Çok değişkenli varyans analizinde gruplar arasındaki farkın tespiti: Bir uygulama, Türkiye Klinikleri Tıp Bilimleri Dergisi, 27(6), 838-845.

Kruger U, Sharma SK, Irwin GW, (2004). Improved Nonlinear Canonical Correlation Analysis Using Genetic Strategies. Control , University of Bath, UK.

Kumar S, Martin EB, Morris J (2002). Non-linear canonical correlation analysis using aRBF network. ESANN'2002 proceedings - European Symposium on Artificial Neural Networks Bruges (Belgium), 24-26 April 2002, Belgium.

Weisberg (2005) Applied Linear Regression. A John Wiley & Sons, Inc. New Jersey USA.

Michailidis G, De Leeuw J (1996). The Gifi System of Descriptive Multivariate Analysis. Technical Report, UCLA Statistics Program, Preprint 204.

Wu A, Hsieh WW (2002). Nonlinear canonical correlation analysis of the tropical Pacific wind stress and sea surface temperature. Climate Dyn, 19, 713–722.

Saraçlı Z, Saraçlı S (2006). Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi öğrencilerinin demografik özellikleri ile üniversite sorunları arasındaki ilişkinin doğrusal olmayan kanonik korelasyon analizi ile incelenmesi. Eskişehir Osmangazi Üniversitesi İİBF Dergisi, 1(1), 27-38.

Süt N, (2001). Doğrusal olmayan kanonik korelasyon analizi ve bir uygulama.Yüksek Lisans Tezi,Trakya Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalı, Edirne.

Theodosiu T, Angelis L, Vakali A (2007). Non-linear correlation of content and metadata information extracted from biomedical article datasets. Department of informatics, School of Natural Sciences, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece.

Türe M, Süt N, Kürüm T, Özbay G (2000). Kardiyolojik verilerin doğrusal olmayan kanonik korelasyon analizi ile incelenmesi. Trakya Tıp Fakültesi Biyoistatistik Anabilim Dalı, Edirne.

Van der Burg, E, De Leeuw J (1988). Use of the multinomial jack-knife and bootstrap in generalized non-linear canonical correlation analysis. Applied stochastic models and data analysis, 4, 159-172.

Van der Burg E, De Leeuw J, Verdegaal R (1984). Non-linear canonical correlation analysis, Leiden: Int Rep RR12–RR84 Department of data theory.

Yeung DS, Tsang ECC (2005). Nonlinear canonical correlation analysis of MRI signals using HDR models. Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, September 1-4.

Yazıcı AC, Ögüş E, Ankaralı H, Gürbüz F (2010). An application of nonlinear canonical correlation analysis on medical data. Turk J. Med. Sci., 40(1).

ÖZGEÇMİŞ

Van'da 1979 yılında doğdu. İlkokulu Van Atatürk İlkokulu'nda, orta öğrenimini Van Atatürk Lisesi'nde tamamladı. Yüksek öğrenimini 2001 yılında Yüzüncü Yıl Üniversitesi Van Sağlık Yüksek Okulu Hemşirelik Bölümünde tamamladı. 2002 yılında Düzce Üniversitesi Tıp Fakültesi'nde hemşire olarak göreve başladı. 2007 yılında Yüzüncü Yıl Üniversitesi, Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalı'nda Yüksek lisans eğitimine başladı. 2006 yılından itibaren de Yüzüncü Yıl Üniversitesi Tıp Fakültesi'nde hemşire olarak çalışmaktadır.