

ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI
ÖLÇME VE DEĞERLENDİRME PROGRAMI

TESTLERİN BOYUTLULUĞUNUN BELİRLENMESİNDE KULLANILAN
YÖNTEMLERDE 1. TİP HATA VE GÜÇ ORANLARININ KARŞILAŞTIRILMASI

DOKTORA TEZİ

GÜL GÜLER

ANKARA, EKİM, 2017

ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI
ÖLÇME VE DEĞERLENDİRME PROGRAMI

TESTLERİN BOYUTLULUĞUNUN BELİRLENMESİNDE KULLANILAN
YÖNTEMLERDE 1. TİP HATA VE GÜÇ ORANLARININ KARŞILAŞTIRILMASI

DOKTORA TEZİ

GÜL GÜLER

DANIŞMAN: PROF. DR. R. NÜKHET DEMİRTAŞLI

ANKARA, EKİM, 2017

Eğitim Bilimleri Enstitüsü Müdürlüğü'ne

Gül GÜLER'in hazırladığı "Testlerin Boyutluluğunun Belirlenmesinde Kullanılan Yöntemlerde 1. Tip Hata ve Güç Oranlarının Karşılaştırılması" başlıklı bu çalışma, jürimiz tarafından Ölçme ve Değerlendirme Anabilim Dalı / Ölçme ve Değerlendirme Programı'nda Doktora Tezi Olarak Kabul Edilmiştir.

İmza

Başkan: Prof. Dr. Ezel TAVŞANCIL



Üye: Prof. Dr. R. Nükhet DEMİRTAŞLI (Danışman)



Üye: Prof. Dr. Duygu ANIL



Üye: Doç. Dr. Fatma MIZIKACI



Üye: Doç. Dr. İsmail KARAKAYA



ONAY

Bu tez Ankara Üniversitesi Lisansüstü Eğitim – Öğretim ve Sınav Yönetmeliği'nin ilgili maddeleri uyarınca yukarıdaki jüri üyeleri tarafından/...../20..... tarihinde uygun görülmüş ve Enstitü Yönetim Kurulunca/...../20..... tarihinde kabul edilmiştir.

Prof. Dr. İsmail GÜVEN
Eğitim Bilimleri Enstitüsü Müdürü

ETİK BİLDİRİM

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.



Gül GÜLER

ÖZET

TESTLERİN BOYUTLULUĞUNUN BELİRLENMESİNDE KULLANILAN YÖNTEMLERDE 1. TİP HATA VE GÜÇ ORANLARININ KARŞILAŞTIRILMASI

Güler, Gül

Doktora, Ölçme ve Değerlendirme Anabilim Dalı

Tez Danışmanı Prof. Dr. R. Nükhet DEMİRTAŞLI

Ekim, 2017, xvi + 89 sayfa

Bu çalışmanın amacı tekboyutlu ve iki boyutlu yapılarda boyutluluk belirlenmesinde kullanılan yöntemlerin 1. tip hata ve güç oranlarının çeşitli koşullar için araştırılması ve her bir koşulun temel etkisinin yanı sıra koşulların ortak etkisini (koşulların etkileşiminin etkisi) karşılaştırmaktır. Bu çalışma kapsamında araştırma sorularına cevap vermek için simülasyon verisinden yararlanılmıştır. Araştırmanın verileri SAS programı kullanılarak üretilmiştir. Veriler, güç çalışması için 2 parametrelili lojistik ve telafisel modelde, iki kategorili (1-0) ve 2 boyutlu yapıya uygun olacak şekilde üretilmiştir. 1. tip hata çalışması için ise 1 parametrelili lojistik modelde, tekboyutlu iki kategorili (1-0) veri üretilmiştir.

Bu araştırmada 1. tip hata çalışması için dağılımın özellikleri (standart normal dağılım ve çarpık dağılım), örneklem büyüklüğü (200, 300, 500, 1000, 2000 ve 3000) ve madde sayısı (10, 20 ve 30) olmak üzere 3 değişken, güç çalışması için ise bu 3 değişkenin yanı sıra boyutlar arası korelasyon (0.25, 0.50, 0.75, 0.90) da dahil olmak üzere 4 değişken manipüle edilmiştir.

Tekboyutluluğun değerlendirilmesinde farklı yöntemlerin performanslarını karşılaştırmak amacıyla hem parametrik hem de parametrik olmayan yöntemlerden yararlanılmıştır. Bu araştırma kapsamında parametrik olmayan yöntemlerden DIMTEST T istatistiği ile Boyutluluk DETECT IDN indeksi kullanılmıştır. Parametrik yöntemlerden ise Doğrusal Olmayan Faktör Analizi (NOHARM) yönteminden yararlanılmıştır.

Araştırma sonucu, standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T istatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için 1. tip hata oranları incelendiğinde, tekboyutluluk kararını vermede en tutarlı sonuçların Doğrusal Olmayan Faktör Analizi yöntemine ait olduğu görülmektedir. Bunun yanı sıra DIMTEST T istatistiği sonuçlarının nispeten daha tutarlı olduğu görülse de kısa testlerde DIMTEST T istatistiğinin boyutluluk belirlemede kullanımının uygun olmayacağı görülmüştür. DETECT IDN indeksinin ise büyük örneklemelerde ve geniş madde havuzlarında kullanımının daha uygun olacağı belirlenmiştir.

Standart normal dağılım gösteren verilere oranla çarpık dağılım gösteren verilerde DETECT IDN indeksi sonuçlarının daha tutarlı olduğu görülse de DIMTEST T istatistiği ve Doğrusal Olmayan Faktör Analizi sonuçlarının ise tekboyutluluk kararını vermede daha isabetli olduğu belirlenmiştir. Güç çalışması sonuçları incelendiğinde ise, standart normal dağılım gösteren verilerde, DIMTEST T istatistiğinin büyük örneklemli koşullarda daha doğru sonuçlar verdiği görülmüştür. Özellikle örneklem büyüklüğünün 300'den daha az olduğu durumlar için DIMTEST T istatistiğinin yanılma oranı belirgin bir şekilde artmaktadır. Aynı zamanda DIMTEST T istatistiğinin çokboyutluluk kararı için, boyutlararası korelasyondan etkilendiği görülmüştür. Boyutluluk DETECT IDN indeksi ve doğrusal olmayan faktör analizi (NOHARM) sonuçlarının tüm koşullarda DIMTEST T istatistiğine göre daha doğru kararlar verdiği görülmektedir. Çarpık dağılım gösteren verilerde ise, testin uzunluğunun 10 maddeden az olduğu koşullarda boyutluluk belirleme yöntemlerinin daha az tutarlı sonuçlar verdiği söylenebilir. Bununla birlikte DETECT IDN indeksi ve Doğrusal olmayan faktör analizi sonuçlarının kendi içinde daha tutarlı olduğu görülmekle birlikte örneklem büyüklüğünün 1000 ve üzeri olduğu koşullarda DIMTEST T istatistiğinin de boyutluluk belirlemede doğru kararlar verdiği görülmüştür.

Anahtar sözcükler: Boyutluluk, Yapı Geçerliği, DIMTEST T İstatistiği, DETECT, Doğrusal Olmayan Faktör Analizi

ABSTRACT**A COMPARISON OF TYPE I ERROR AND POWER RATES IN PROCEDURES
USED DETERMINING TEST DIMENSIONALITY**

Güler, Gül

Dissertation, Department of Educational Measurement and Evaluation

Advisor: Prof. Dr. R. Nükhet DEMİRTAŞLI

October, 2017, xvi + 89 pages

The aim of this study is to examine type 1 error and power rates of the methods used to determine dimensionality in one-dimensional and two-dimensional structures under various conditions and to compare the common effect of the conditions (the effect of interaction of the conditions) in addition to the main effect of each condition. In this study, simulation data were used to answer research questions. The data of the study were generated using the SAS. The data were generated in a 2-parameter logistic and compensatory model for power study in a way to be suitable for two-categorical (1-0) 2-dimensional structures. For type 1 error analysis, unidimensional two-categorical (1-0) data were generated in a 1-parameter logistic model.

In this study, three variables were manipulated for type 1 error analysis as properties of distribution (standard normal distribution and skewed distribution), sample size (200, 300, 500, 1000, 2000 and 3000), and test length (10, 20 and 30), and four variables for power study including correlations between dimensions (0.25, 0.50, 0.75, 0.90) in addition to these three variables.

Both parametric and nonparametric methods were used to compare the performance of different methods in evaluating unidimensionality. In this research, the nonparametric methods included DIMTEST T statistics and Dimensionality DETECT IDN index. As for the parametric methods, Non-Linear Factor Analysis (NOHARM) was used.

Analysis of type 1 error for DIMTEST T statistics, Dimensionality DETECT IDN index and Nonlinear Factor Analysis in data with standard normal distribution showed that the most consistent results were yielded by Nonlinear Factor Analysis method in determining unidimensionality for different sample sizes and different item numbers. Besides, though the results of DIMTEST T statistics were relatively more consistent, it is considered that DIMTEST T statistics would not be useful to determine dimensionality in short tests. Apart from that, the use of DETECT IDN index would be more appropriate in large samples and large item pools.

Although the results of the DETECT IDN index were consistent in the case of the data with skewed distribution compared to the standard normal distribution, it can be said that the results of DIMTEST T statistics and Nonlinear Factor Analysis were more accurate in determining unidimensionality. Power study results implied that the DIMTEST T statistics provided more accurate results for large samples in the case of data with standard normal distribution. Particularly, for the sample size less than 300, the error rate of DIMTEST T statistics increased visibly. At the same time, it can be said that the DIMTEST T statistics were affected by the correlation between dimensions for the decision of multidimensionality. It seems that Dimensionality DETECT IDN index and nonlinear factor analysis (NOHARM) results are apt to make more accurate decisions than the DIMTEST T statistics in all conditions. On the other hand, in the case of the skewed data, it can be said that dimensionality determination methods yield less consistent results when the length of the test was less than 10 items. Also DETECT IDN index and nonlinear factor analysis results were found to have higher self-consistence; however, DIMTEST T statistics were found to make the right decision in determining dimensionality when the sample size is 1000 and above.

Keywords: Dimensionality, Construct Validity, DIMTEST T Statistics, DETECT IDN index, Nonlinear Factor Analysis

ÖNSÖZ

Eğitim ve psikolojide kullanılan testlere bakıldığında birçoğunun birden çok örtük özelliği ölçtüğü görülmektedir. Örneğin fen alanında bir test bir bilimsel süreç becerilerini ölçmek üzere geliştirilmişken aynı zamanda okuduğunu anlama becerisini de ölçebilir. Bu nedenle ölçülmek istenen yapının tekboyutlu mu yoksa çok boyutlu mu olduğunu bilmekte yarar vardır. Bu durum testin oluşturulma ve uygulanma amacı göz önünde bulundurulduğunda test puanlarına dayalı olarak bireyler hakkında verilen kararların geçerliğini etkileyecektir. Bu tez çalışmasında da parametrik ve parametrik olmayan boyutluluk belirleme yöntemlerin çeşitli koşullara göre etkililiği araştırılmıştır.

Doktora eğitimim süresince, her zaman bana destek olan, fikirleri ve önerileriyle bana ışık tutan değerli hocam ve danışmanım Prof. Dr. R. Nükhet DEMİRTAŞLI'ya sonsuz teşekkürlerimi sunarım.

Tez izleme komitesinde yer alan ve değerli bilgileriyle tezime katkıda bulunan saygıdeğer hocalarım Prof. Dr. Ezel TAVŞANCIL ve Doç. Dr. Fatma MIZIKACI'ya; tez jürimde yer alarak değerli görüş ve önerilerini esirgemeyen, tezimin son halini almasında katkıları bulunan Prof. Dr. Duygu ANIL ve Doç. Dr. İsmail KARAKAYA'ya teşekkürlerimi sunarım.

Lisansüstü eğitimimin başından beri bana dost, kardeş olan aynı zamanda doktora eğitimim boyunca Ankara'da bana evini açıp kendi evimmiş gibi hissettiren ve tüm bunların yanı sıra akademik hayatım boyunca beni yüreklendiren, tezimin her aşamasıyla ilgilenip zor zamanlarımda bana olan inancıyla güç veren, dostum Dr. Sakine GÖÇER ŞAHİN'e sonsuz teşekkürlerimi sunarım.

Son olarak bu zor süreçte beni anlayan, koşulsuz bana destek olan sevgili eşim Deniz GÜLER'e, bugünlere gelmemde büyük emeği olan babam ve biricik kardeşlerime, hiçbir zaman desteklerini esirgemeyen, yükümü hafifleten kayınvalidem ve kayınpederime çok teşekkür ederim.

Gül GÜLER



Oğlum Can Deniz'e

İÇİNDEKİLER

ONAY.....	ii
ETİK BİLDİRİM.....	iii
ÖZET.....	iv
ABSTRACT	vi
ÖNSÖZ.....	viii
İÇİNDEKİLER.....	x
ÇİZELGELER TABLOSU	xiii
KISALTMALAR	xvi
BÖLÜM I	1
1. GİRİŞ.....	1
1.1. Problem Durumu.....	1
1.1.1.Yapı Geçerliği	1
1.1.2. Boyutluluk.....	3
1.1.3. Madde Tepki Kuramı	4
1.1.3.1.Tekboyutluluk.....	5
1.1.3.2.Yerel bağımsızlık.....	6
1.1.3.3. Tekboyutlu Madde Tepki Kuramı Modelleri	6
1.1.3.3.1.Bir Parametrelili Lojistik Model ve Rasch Modeli	6
1.1.3.3.2. İki Parametrelili Lojistik Model.....	7
1.1.3.3.3. Üç Parametrelili Lojistik Model.....	8
1.1.4. Çok Boyutlu Madde Tepki Kuramı.....	8
1.1.4.1.Çok Boyutlu Madde Tepki Kuramı Modelleri	9
1.1.5. Boyutluluk Belirleme Yöntemleri.....	11
1.1.5.1. Parametrik Olmayan Yöntemler (Non-parametric Methods).....	13
1.1.5.1.1. Boyutluluk Testi (DIMTEST Süreci)	13

1.1.5.1.2. DETECT İndeksi (Dimensionality Evaluation to Enumerate Contributing Traits)	14
1.1.5.2. Parametrik Yöntemler (Parametric Methods).....	15
1.1.5.2.1. Açımlayıcı Faktör Analizi.....	15
1.1.5.2.2. Doğrusal Olmayan Faktör Analizi: Normal Ogive Harmonic Analysis Robust Method (NOHARM)	16
1.1.5.2.2.1. Tanaka Uyum İyiliği İndeksi (TIGF).....	17
1.1.5.2.2.2. Artıkların Ortalama Kareler Karekökü (RMSR).....	17
1.1.5.2.2.3. Yaklaşık Ki-Kare ($\chi^2_{G/D}$).....	18
1.2. İlgili Araştırmalar	18
1.2.1. Boyutluluk Yöntemlerin Karşılaştırıldığı Çalışmalar	19
1.2.2. İlgili Yöntemlerin Kullanıldığı Ulusal Çalışmalar.....	23
1.3. Amaç.....	24
1.4. Önem.....	25
1.5. Sınırlılıklar	26
BÖLÜM II.....	27
2. YÖNTEM.....	27
2.1. Araştırmanın Modeli.....	27
2.2. Veri Üretim Çalışması	27
2.3.Verilerin Analizi	33
BÖLÜM III.....	35
3. BULGULAR ve YORUMLAR.....	35
3.1. "1-0 Puanlanan Testlerde Çeşitli Boyutluluk Belirleme Yöntemlerine Göre 1. Tip Hata Oranı Nasıl Değişmektedir?" Bir Başka İfadeyle 1-0 Puanlanan Testlerde Çeşitli Boyutluluk Belirleme Yöntemlerine Göre Testin Uzunluğunun, Dağılımın Özelliğinin ve Örneklem Büyüklüğünün Manipule Edildiği Tekboyutlu Verilerden Elde Edilen 1. Tip Hata Oranları Nasıl Değişmektedir?	35

3.2. “1-0 Puanlanan Testlerde Çeşitli Boyutluluk Belirleme Yöntemlerine Göre Testin Gücü Nasıl Değişmektedir?” Bir Başka İfadeyle 1-0 Puanlanan Testlerde Çeşitli Boyutluluk Belirleme Yöntemlerine Göre Testin Uzunluğunun, Boyutlararası Korelasyon Derecesinin, Dağılımların ve Örneklem Büyüklüğünün Manipule Edildiği İki Boyutlu Veriler İçin Testin Gücü Nasıl Değişmektedir?	46
3.3. “1-0 Puanlanan Testlerde Çeşitli Boyutluluk Belirleme Yöntemlerine Göre Standart Normal ve Çarpık Dağılım Gösteren Verilerde 1. Tip Hata Oranı Ve Testin Güç Oranları Nasıldır?”	64
BÖLÜM IV	71
4. SONUÇ ve ÖNERİLER	71
4.1. Sonuçlar	71
4.2. Öneriler	72
EKLER	76
EK 1. Güç Çalışması için Veri Üretmede Kullanılan SAS Kodları	77
EK 2. 1. Tip Hata Çalışması için Veri Üretmede Kullanılan SAS Kodları	79
EK 3. Testin Gücü için Yöntemlere İlişkin Programların Çalıştırılmasında Kullanılan Örnek SAS Kodları	81
EK 4. 1. Tip Hata Çalışması için Yöntemlere İlişkin Programların Çalıştırılmasında Kullanılan Örnek SAS Kodları.....	82
KAYNAKÇA	83
ÖZGEÇMİŞ.....	89

ÇİZELGELER TABLOSU

1. Boyutluluk belirlemede I. Tip hata ve testin gücüne ilişkin istatistiksel karşılaştırma	3
2. İkili puanlanan maddelerden oluşan testlerin boyutluluğunu belirlemede kullanılan parametrik ve parametrik olmayan yöntemler ve bilgisayar yazılımları	12
3. Veri üretiminde kullanılan değişken ve koşulları.....	29
4. 1. Tip hata için araştırma deseni.....	29
5. Güç çalışması için araştırma deseni	30
6. Madde parametreleri.....	32
7. Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T İstatistiği için 1. tip hata oranları.....	35
8. Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre Boyutluluk DETECT IDN indeksi için 1. tip hata oranları	37
9. Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre Doğrusal Olmayan Faktör Analizi için 1. tip hata oranları.	38
10. Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T İstatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için 1. tip hata oranları	39
11. Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T İstatistiği için 1. tip hata oranları	41
12. Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre Boyutluluk DETECT IDN indeksi için 1. tip hata oranları	43
13. Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre Doğrusal Olmayan Faktör Analizi için 1. tip hata oranları.....	44
14. Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T İstatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için 1. tip hata oranları	45

15. Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre DIMTEST T İstatistiği için güç oranları 47
16. Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre Boyutluluk DETECT IDN İndeksi için güç oranları..... 50
17. Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre Doğrusal Olmayan Faktör Analizi için güç oranları 52
18. Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre DIMTEST T İstatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için güç oranları 54
19. Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre DIMTEST T İstatistiği için güç oranları 57
20. Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre Boyutluluk DETECT IDN İndeksi için güç oranları 59
21. Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre Doğrusal Olmayan Faktör Analizi (NOHARM) için güç oranları 61
22. Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre DIMTEST T İstatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için güç oranları 63
23. Standart normal dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre DIMTEST T istatistiği için 1. Tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon derecesine göre DIMTEST T istatistiği için testin gücü oranları..... 65

24. Standart normal dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Boyutluluk DETECT IDN indeksi için 1. Tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon derecesine göre Boyutluluk DETECT IDN indeksi için testin gücü oranları..... 66
25. Standart normal dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Doğrusal Olmayan Faktör Analizi için 1. Tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon derecesine göre Doğrusal Olmayan Faktör Analizi için testin gücü oranları..... 67
26. Çarpık dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre DIMTEST T istatistiği için 1. Tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon derecesine göre DIMTEST T istatistiği için testin gücü oranları..... 68
27. Çarpık dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Boyutluluk DETECT IDN indeksi için 1. Tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon derecesine göre Boyutluluk DETECT IDN indeksi için testin gücü oranları 69
28. Çarpık dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Doğrusal Olmayan Faktör Analizi için 1. Tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon derecesine göre Boyutluluk Doğrusal Olmayan Faktör Analizi için testin gücü oranları..... 70

KISALTMALAR

AT1, AT2	Değerlendirme Alt testleri (Assessment Subtest)
ÇBMTK	Çok Boyutlu Madde Tepki Kuramı
DFA	Doğrulayıcı Faktör Analizi
FA	Faktör Analizi
PT	Bölümleme Alt Testi (Partitioning Subtest)
MTK	Madde Tepki Kuramı



BÖLÜM I

1. GİRİŞ

Bu bölümde araştırmaya ilişkin problem durumu ele alınmış; araştırmanın amacına, önemine ve sınırlılıklarına yer verilmiştir.

1.1. Problem Durumu

Geçerlik kanıtları sunmak, eğitim ve psikolojide yer alan sınavların kullanımı açısından önemlidir. Geçerlik kanıtlarına sadece sınavlardan elde edilen puanlara anlamlı bir zemin hazırlamak için değil aynı zamanda puanın kullanımının doğurduğu sosyal sonuçları bilmek için de ihtiyaç duyulur (Messick, 1995). Eğitim ve psikolojide test ve ölçek geliştirme sürecinde boyutluluk kavramı geçerlik çalışmalarında sıklıkla kullanılmaktadır. Boyutluluk bir testte yer alan maddeler ile testin ölçtüğü düşünülen örtük özellik arasındaki ilişki şeklinde tanımlanabilir (Svetina, 2011). Bir başka ifadeyle, boyutluluk, testteki maddelerin, hangi boyutlarla ilişkili olduğunu ifade etmektedir. Boyutluluk, bir testin ya da madde setinin ölçtüğü becerilerin veya yapıların sayısı ile ilgilidir. Boyutluluk belirleme süreci, ölçme modelinin tek boyutlu veya çokboyutlu olduğuna bakılmaksızın göz önünde bulundurulması gereken önemli bir konudur (Embretson ve Reise, 2000). Bir test kuramsal bir yapıya sahip olup özel bir amaç için hazırlanır. Testin altında yatan yapı incelenerek doğrulanmalıdır. Bu bağlamda yapı geçerliliği çalışmaları, eğitim ve psikolojide araçların kullanımı açısından önemlidir ve bununla birlikte test ve ölçeklerin boyutluluğunun değerlendirilmesinde gerekli adımlardan biridir.

1.1.1.Yapı Geçerliği

Bireylere atfedilen psikolojik özellikler bir yapı olarak kabul edilir ve bu yapılar doğrudan gözlenemediği için bireylerin bir ölçme aracına verdiği tepkiler doğrultusunda ortaya konabilir (Yüce, 2012). Yapı geçerliliği, bir testin ölçülmek istenen kuramsal

yapıyı ortaya koyabilme derecesidir (Anastasi ve Urbana, 1997). Ölçülmek istenen bir özellik doğası gereği birden fazla örtük özellikle ilişkili olabilir. Eğitim ve psikolojide kullanılan testlere bakıldığında birçoğunun birden çok örtük özelliği ölçtüğü görülmektedir. Örneğin fen alanında bir test bir bilimsel süreç becerilerini ölçmek üzere geliştirilmişken aynı zamanda okuduğunu anlama becerisini de ölçebilir. Bu nedenle ölçülmek istenen yapının tekboyutlu mu yoksa çok boyutlu mu olduğunu bilmekte yarar vardır. Bu durum testin oluşturulma ve uygulanma amacı göz önünde bulundurulduğunda test puanlarına dayalı olarak bireyler hakkında verilen kararların geçerliğini etkileyecektir. Bir testteki maddelerin boyutluluğunun belirlenmesi, verilerin istatistikî analizlerini de şekillendireceği için son derece önemlidir (Hambleton, 1991; Svetina, 2011; Zhang, 2007).

Çeşitli araştırmalar, özellikle geniş ölçekli testlerin geliştirilmesi, değerlendirilmesi ve sürdürülmesinde testlerin boyutluluğunun değerlendirilmesiyle geçerliğe kanıt sağlamanın önemine değinmiştir (Hattie, 1985; Yeh, 2007).

Boyutluluk değerlendirmesi yapılması, testi geliştiren tarafından tanımlanan yapının tespit edilmesini ve testin altında yatan yapıyı ne kadar ölçebildiğinin anlaşılmasını kolaylaştırır (Mroch ve Bolt, 2006; Yeh, 2007). Bir diğer deyişle, testi hazırlayan bireyler, hangi alanların ölçüldüğünü ve ölçülen alanlar arasındaki ilişkileri ortaya çıkarmak istiyorsa boyutluluk değerlendirmesi yapar.

Verilerin psikometrik modellemesini yaparak araştırmacılar öğrenciler ile ilgili çıkarımda bulunacak dayanaklara ulaşır. Bu tür çıkarımlar yapabilmek için analizde kullanılan psikometrik modeller teknik açıdan sağlam olmalı ve testlerden alınan verilerle uyumlu hale getirilmelidir (Svetina, 2011).

Yen (2007), test boyutluluğu değerlendirmesinin (1) ölçülen boyutları doğrulamak; (2) boyutlar arasındaki ilişkileri anlamak; (3) zaman içerisinde test yapısını gruplara göre incelemek ve sürdürmek açısından önemini vurgulamıştır. Buna ek olarak boyutluluğun, yapı geçerliliği, yapının yeterince temsil edilmemesi gibi tehditlerin tespit edilmesinde de işe yarayacağını ifade etmiştir.

Bir ölçme işlemi aslında çok boyutlu olup da tekboyutlu gibi işlem görürse test puanlarının yorumlanması zarar görecektir ve yapılan ölçme işlemlerinin geçerliği zedeleneyecektir (Göçer Şahin, 2016; Tournon, Lizasoain ve Joaristi, 2012). Boyutluluğun

belirlenmesi, bunun yanı sıra tek boyutluluğun ne derece ihmal edildiğinin belirlenmesi ve I. tip hata ile testin gücünün ortaya konması uygulanan testler sonucu verilen kararların geçerliği açısından önemlidir. Boyutluluk belirlemede 1.tip hata ve testin gücüne ilişkin istatistiksel karşılaştırma Çizelge 1’de verilmiştir:

Çizelge 1.

Boyutluluk belirlemede I. tip hata ve testin gücüne ilişkin istatistiksel karşılaştırma

		GERÇEK DURUM	
		Ho: Doğru	Ho: Yanlış
İSTATİSTİKSEL KARAR	*Ho: Kabul	Doğru Karar ($1-\alpha$) (Doğru Ho hipotezini kabul etme olasılığı)	II. Tip Hata (β) (Yanlış Ho hipotezini kabul etme olasılığı;)
	Ho: Red	I. Tip Hata (α) (Doğru Ho hipotezini reddetme olasılığı)	Doğru Karar (Güç)($1-\beta$) (Yanlış Ho hipotezini reddetme olasılığı)

*Ho: Test tek boyutludur ($d=1$)

**H₁: Test çokboyutludur ($d>1$)

Çizelge 1’e göre, bir test tekboyutlu iken yani gerçekte Ho hipotezi doğru iken istatistiksel kararlar H₁ hipotezinin kabul edilmesi yani test çok boyutludur denmesi I. tip hataya neden olmaktadır. Bir test çok boyutlu iken Ho hipotezinin kabul edilmesi bir başka ifadeyle tekboyutludur denmesi ise II. tip hataya neden olmaktadır. Tüm bunların yanı sıra bir test gerçekte çok boyutlu iken istatistiksel kararlar da çok boyutludur denmesi ise testin gücünü göstermektedir. Bu nedenle bütün bu durumların tespit edilmesi verilecek kararların geçerliliği ile doğrudan ilişkili olup test etmenin gerekli olduğu düşünülmektedir.

1.1.2. Boyutluluk

Boyutluluk çeşitli biçimlerde tanımlanmaktadır. Boyutluluk, bir dizi test madde tepkilerine temel oluşturan ve maddeler arasındaki varyans-kovaryansların istatistiksel açıklamasını yapan özelliklerin sayısı olarak ifade edilmektedir (Hattie ve ark., 1996;

Stout ve ark., 2001; Zhang, 2007). Bir testin boyutluluğunun belirlenmesi iki çeşit probleme ulaştırmaktadır (Touren, Lizasoain ve Joarasti, 2012);

1. Tek boyutluluğu doğrulama ya da aksini ispatlama
2. Eğer gerekirse testin çokboyutlu yapısının tanımı

Boyutluluğun değerlendirmesinin ihmal edilmesi madde tepki kuramı ve klasik test kuramı çerçevesinde yanlış uygulamalara neden olabilir. Klasik test kuramı çerçevesinden düşünüldüğünde tekboyutluluk temel varsayım olarak ifade edilmese de aslında madde puanlarının toplanabilirliği açısından son derece önemlidir (McDonald, 1999; Yen, 2007). Testlerin boyutluluğunun belirlenmesinde bir diğer önemli nokta ise, bir testteki boyutlar ayırt edilebilir olduğu zaman alt puanları yorumlamak uygun olur ve eğer başat tek bir boyut varsa tek bir toplam puan raporlanır (Haladyna, 2004). Eğer bir testte birden fazla boyut açıkça ifade edilmişse bu boyutlar için ayrı ayrı alt puanları raporlamak uygun olacaktır. Boyutluluk değerlendirmesi sonucu yapı tek bir boyut gösteriyorsa tek bir toplam puan raporlaması tercih edilebilir. Boyutluluğun değerlendirmesi, alt puanların doğru yorumlanmasına ve alt puanların öğretim faaliyetlerinde kullanılmasına yönelik kanıt toplamaya yardımcı olabilir (Yen, 2007).

Madde tepki kuramı çerçevesinden de düşünüldüğünde, kuramın temel varsayımlarından biri olan tekboyutluluğun göz ardı edilmesi, çok boyutlu bir verinin tekboyutluymuş gibi düşünülmesi yanlış madde ve birey parametrelerinin kestirimine, test eşitlemelerine, madde yanlılığına, test oluşturma ve puanların raporlaştırılmasına neden olacaktır. Boyutların geçerliği açısından bakılacak olursa testi hazırlayan kişi, gruplar genelinde puanların karşılaştırılabilirliğini dikkatli değerlendirmelidir. Bu nedenle eğitimde ve psikolojide kullanılan bir ölçme aracının boyutluluğunun araştırılması tekboyutlu MTK modelleri uygulamaları için son derece önemlidir.

1.1.3. Madde Tepki Kuramı

Madde tepki kuramı, bir bireyin testle ölçülen yetenek düzeyini kestirmek için bir temel oluşturur. Kurama göre bir bireyin yetenek düzeyi, bireyin maddelere verdiği tepkilerden kestirilir. KTK'da madde ve testin özellikleri gruba bağlıdır. Bu nedenle farklı test veya madde ile karşılaşıldığında grubun özelliği de değişmektedir. Bu nedenle KTK'nın, test geliştirme, test eşitleme, yansız parametre kestiriminde yetersiz

kaldığı söylenebilir (Embretson ve Reise, 2000). Tekboyutluluk ve yerel bağımsızlık MTK'nın iki temel varsayımı olup boyutluluk belirleme sürecinde son derece önemlidir.

1.1.3.1. Tekboyutluluk

Tekboyutluluk, bir testte yer alan her bir maddenin tek bir yeteneği ölçmesi olarak ifade edilmektedir (Nandakumar ve Stout, 1993). Bir ölçme aracının tekboyutluluğu, tek bir örtük özelliği ölçen maddeler arasındaki istatistiksel bağımlılığı ifade etmektedir (Crocker ve Algina, 1986).

Bir testte yer alan maddelerin doğru cevap olasılıklarının belirlenmesinde birden fazla yetenek boyutu etkili olmamalıdır. Testte yer alan herhangi bir maddenin doğru cevaplanma olasılığında baskın olarak iki veya daha fazla özellik etkinse, tekboyutluluk varsayımı karşılanamayacaktır (Sünbül, 2011). Eğer tekboyut varsa her bir öğrenci için doğru cevaplanan maddelerin toplamı alınarak işlemlerin yapılması genelde tercih edilir. Fakat birden fazla boyut varsa bu uygulama yeterli olmaz. Bu nedenle yapılması gereken örtük boyutlarda sunulan maddelerin çeşitli setleri için farklı puanların kestirilmesidir (Touren, Lizasoain ve Joarasti, 2012).

Tekboyutluluk varsayımının karşılanıp karşılanmadığını belirlemek için sıklıkla, elde edilmiş olan ham verilere faktör analizi yapılmakta ve başat bir faktör arayışına girilmektedir. Başat bir faktör bulunduğunda tekboyutluluk varsayımının karşılandığı ifade edilerek, madde ve yetenek parametrelerinin kestirimleri yapılmaktadır (Sünbül, 2011). Boyutluluk ve bununla birlikte tekboyutluluk incelemelerinde faktör analizi sıkça kullanılmasına rağmen birçok araştırmada da sadece faktör analizi ile tekboyutluluk incelenmesinin yeterli olmadığı ifade edilmiş olup bu incelemeler için paralel analiz, iki faktörlü çözümleme (bifactor analysis), boyutluluk testi (DIMTEST), boyutluluk DETECT indeksi, doğrusal olmayan faktör analizi, hiyerarşik kümeleme analizi yöntemi gibi çok sayıda başka yöntemler de önerilmektedir (Finch ve Monahan, 2008; Ledesma ve Valero-Mora, 2007; Özbek, 2012; Özkan ve Güvendir, 2014; Reichenberg, 2013; Sünbül, ve Seo, 2012; Svetina, 2011; Svetina ve Levy, 2014; Touren, Lizasoain ve Joarasti, 2012; Yen, 2007).

Tekboyutluluk, daha net bir tanım yapılmaya kadar homojenlik ve iç tutarlılık kavramlarıyla aynı anlamda kullanılırken, Örtük Özellik Modelleri'ne (şimdiki adıyla Madde Tepki Kuramı-MTK modelleri) ilginin artmasıyla daha açık bir şekilde ifade

edilmeye başlanmıştır (Hattie, 1985). Homojenlik daha çok madde korelasyonlarının benzerliğine dayanırken; tekboyutluluk, tek bir özelliğin ölçülmesi olarak bilinmektedir (Hattie, 1985).

Tekboyutluluk varsayımının karşılanması MTK'nın önemli ikinci varsayımı olan yerel bağımsızlık varsayımının sağlanmasına da önemli ölçüde katkı sunacaktır.

1.1.3.2. Yerel bağımsızlık

Yerel bağımsızlık, bireyin bir maddeye verdiği tepkinin diğer maddelere verdiği tepkilerden istatistiksel olarak bağımsız olmasıdır (Reckase, 2009). Tekboyutluluk varsayımının karşılanmaması yerel bağımsızlık varsayımını da beraberinde zedeleyecektir. Eğer maddeye verilecek olan tepki birden fazla yeteneğe bağlıysa, maddeleri bağımsız yapmak için süreçte etkili olan yeteneklerden birini sabitlemek yeterli olmayacaktır ve sabitlenmemiş olan diğer yetenek, maddeler arasında bağımlılık oluşturacaktır (Sünbül, 2011).

1.1.3.3. Tekboyutlu Madde Tepki Kuramı Modelleri

Madde tepki kuramı modeli, bireyin, bir özelliğe ilişkin yetenek düzeyi (θ_i) ile testteki performansı arasındaki ilişkinin matematiksel olarak ifade edilmesidir. Yetenek ile performans arasındaki ilişkiyi açıklayan çok sayıda MTK modeli vardır (Hambleton ve Swaminathan, 1985).

MTK modelleri, iki kategorili ve çok kategorili olmak üzere iki şekilde ele alınmaktadır (Hambleton ve Swaminathan, 1985). Alanyazında, iki kategorili puanlanan veri için lojistik ve ogive modeller önerilmiştir ancak iki kategorili puanlanan veri için kullanılan normal ogive modeller integral içeren bir formüle dayanması nedeniyle daha güçtür ve onun yerine lojistik modeller geliştirilmiştir (Hambleton ve Swaminathan, 1985). Bu araştırmada lojistik model kullanıldığından lojistik modellere değinilmiştir.

1.1.3.3.1. Bir Parametrelili Lojistik Model ve Rasch Modeli

Bir parametrelili lojistik model, iki ve üç parametrelili lojistik modellerin özel bir hali olup, bu modelle kestirilmeye çalışılan sadece güçlük parametresidir (Hambleton ve

Swaminathan, 1985). Madde güçlük parametresine bağlı olarak belirli bir yetenek düzeyindeki bir kişinin maddeyi doğru cevaplama olasılığına ilişkin fonksiyon belirlenir. Rasch modeli de bir parametrelili lojistik model gibi bir parametrelidir ancak aralarındaki tek fark Rasch modelinde a parametresi 1 kabul edilirken, bir parametrelili lojistik modelde a değerinin ortalaması alınarak sabitleştirilmektedir. Bu model, şans parametresinin bulunmadığı ya da ihmal edilebilecek kadar küçük olduğu varsayımına dayalıdır. Bu modelde bir s bireyinin i maddesini doğru cevaplama olasılığı aşağıdaki gibidir:

$$P(X_{is} = 1 | \theta_s, \beta_i) = \frac{\exp(\alpha(\theta_s - \beta_i))}{1 + \exp(\alpha(\theta_s - \beta_i))} \quad (1)$$

X_{is} = s bireyin i maddesine verdiği cevap (1 veya 0)

θ_s = s bireyinin yetenek düzeyi

β_i = i maddesinin güçlüğü

α_i = i maddesinin ayırtıcılığı

1.1.3.3.2. İki Parametrelili Lojistik Model

Bu modelde, bir parametrelili modelde yer alan b_i güçlük parametresine ek olarak α_i parametresi de hesaba katılır (Hambleton ve Swaminathan, 1985). Madde karakteristik eğrisinin bükülme noktasındaki eğim olarak belirlenen bu parametre madde karakteristik eğrileri farklı olan maddeler için farklı eğimlere sahiptir. Bu modele göre bir s bireyinin i maddesini doğru cevaplama olasılığı aşağıdaki gibidir:

$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i) = \frac{\exp[\alpha_i(\theta_s - \beta_i)]}{1 + \exp[\alpha_i(\theta_s - \beta_i)]} \quad (2)$$

X_{is} = s bireyin i maddesine verdiği cevap (1 veya 0)

θ_s = s bireyinin yetenek düzeyi

β_i = i maddesinin güçlüğü

α_i = i maddesinin ayırtıcılığı

1.1.3.3. Üç Parametrelili Lojistik Model

Bu model, iki parametrelili lojistik modeldeki ayırtedicilik ve güçlük parametrelerinin yanı sıra maddenin ölçtüğü özelliğe sahip olmayanların şansla cevap verebilme olasılığının gözönünde bulundurulduğu c parametresinin eklenmesiyle oluşan modeldir. Bir ve iki parametrelili modelde bu değer sıfır kabul edildiği için madde karakteristik eğrisi y eksenini sıfır noktasında kesmektedir. Üç parametrelili lojistik modele ilişkin olasılık fonksiyonu aşağıdaki gibidir:

$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i, \gamma_i) = \gamma_i + (1 - \gamma_i) \frac{\exp[\alpha_i(\theta_s - \beta_i)]}{1 + \exp[\alpha_i(\theta_s - \beta_i)]} \quad (3)$$

X_{is} = s bireyinin i maddesine verdiği cevap (1 veya 0)

θ_s = s bireyinin yetenek düzeyi

β_i = i maddesinin güçlüğü

α_i = i maddesinin ayırtıcılığı

γ_i = i maddesinin en düşük asimptotu (şans parametresi)

Geliştirilen bir test veya ölçek uygulamalarında en sık karşılaşılan durumlardan biri, kullanılan ölçme araçlarının tek boyutlu olmamasıdır. Bu durumda MTK'nın temel varsayımı olan tekboyutluluğun ihlal edilmesine neden olmaktadır (Çakıcı Eser, 2015). Bunun yanı sıra çok boyutlu veri grubu için tek bir örtük özellik boyutunun olduğunu varsaymak, MTK'nın değişmezlik özelliğini de zedelemektedir (Ackerman, 1994). Bu durumda tekboyutlu MTK modelleri yerine çok boyutlu veri için geliştirilmiş olan çok boyutlu madde tepki kuramının (ÇBMTK) kullanılmasını beraberinde getirmektedir.

1.1.4. Çok Boyutlu Madde Tepki Kuramı

Bir testin yanıtlanmasından sorumlu birden fazla boyut varsa örtük değişken modeli çok boyutlu olarak ifade edilmektedir. Daha önce de ifade edildiği gibi bir testteki maddelerin yanıtlanmasında birden fazla örtük özellik etkili olabilir. Örneğin okuduğunu anlama becerisi, madde formatı, dış müdahale, paragraf bağımlılığı, hız testi gibi nedenler tek boyutluluğu zedelediği gibi aynı zamanda yerel bağımlılığa da neden olmaktadır (Yen, 1993). Eğitimde başarının ölçüldüğü testlerin çoğunun çok boyutlu

yapıda olduğu söylenebilir. Çünkü başarıda bellek, dikkat, okuduğunu anlama, karar verme gibi pek çok bilişsel süreç birbirini tamamlar.

Çok boyutlu Madde Tepki Kuramı, yapı veya boyut olarak ifade edilen iki veya daha fazla örtük değişken ile testi alanın belirli bir test maddesini doğru olarak cevaplama olasılığı arasındaki ilişkiyi matematiksel bir model ile açıklamaktadır (Köse, 2012). Çok boyutlu MTK, tekboyutlu MTK modellerinin çok boyutluluğa uyarlanması olarak da düşünülebilir (Ackerman, Gierl ve Walker, 2003). Kuramda, tekboyutluluk varsayımının göz ardı edilmesi söz konusudur. Tekboyutlu MTK modellerinde doğru cevaplama olasılığı tek bir yeteneğe bağlı iken, çok boyutlu MTK’da doğru cevaplama olasılığı ölçülmeye çalışılan çok sayıda özelliğin bir fonksiyonudur.

Örtük bir veriye çok boyutlu MTK’yı uygulamak için veri setinin çok boyutlu olmasının yanı sıra yerel bağımsızlık ve monotonluk varsayımlarının da karşılanması gerekmektedir. Yerel bağımsızlık, tekboyutlu MTK’da olduğu gibi bireyin bir maddeye verdiği tepkinin diğer maddelere verdiği tepkilerden istatistiksel olarak bağımsız olmasıdır (Reckase, 2009). Monotonluk varsayımı ise maddenin ilişkili olduğu yeteneklerden herhangi birinin artmasıyla bireyin maddeyi doğru cevaplama olasılığının da artması durumudur.

1.1.4.1.Çok Boyutlu Madde Tepki Kuramı Modelleri

Çok boyutlu MTK modeli, θ örtük özellik vektörünün verdiği bilgiye bağlı olarak iki temel grupta sınıflandırılabilir: Telafisel (compensatory) model ve telafisel olmayan (noncompensatory) modeldir. Telafisel model, θ koordinatlarının doğrusal bir kombinasyonuna dayanmaktadır. Bu kombinasyonlar, çeşitli θ değerlerinin kombinasyonlarının toplamını ifade etmektedir. Telafisel modeller, doğru cevaplama olasılığı açısından bir boyuttaki yüksek yeteneğin diğer boyuttaki düşük yeteneği telafi etmesine izin verdiği modeller olarak tanımlanır (Ansley ve Forsyth, 1985; akt. Spray, Dawey, Reckase, Ackerman ve Carison, 1990). Eğer bir θ değeri düşük ve bir diğer θ değeri yeteri kadar yüksek ise toplam yine aynı olacaktır. Telafisel model çok boyutlu yapılarda bir yetenek düzeyindeki yetkinliğin diğer boyut ya da boyutlardaki eksik olan yetkinliği veya beceriyi tamamlaması durumunda kullanılabilecek olan bir modeldir. Bir başka ifadeyle bir θ değerinin diğerini telafi etmesi, dengelemesi söz konusu

olduğundan bu tür modellere telafisel (compensatory) modeller adı verilmektedir (Reckase, 2009).

Telafisel modelde, bir maddeyi doğru yanıtlamak için gerekli olan en az iki boyuttan birinin yeterli olmadığı durumda, diğer boyut veya boyutlardaki özellik aracılığı ile maddenin doğru yanıtlanması söz konusudur. Örneğin, Yabancı Dil Sınavı'nda (YDS) astım hastalığı ile ilgili bir konu hakkında bilgi veren bir maddeyi düşünelim. Böyle bir maddede hem astım hastalığı bilgisi hem de yabancı dil alanında okuduğunu anlama becerisi yoklanmaktadır. Ancak astım hastalığı bilgisi yüksek olan bir birey yabancı dile ilişkin yeteneği düşük olsa da maddeyi doğru cevaplayabilir ya da bireyin astım hastalığı ile ilgili bilgisi yeterli olmasa bile sahip olduğu yüksek yabancı dil bilgisi ile maddeyi doğru cevaplaması söz konusu olabilir. Burada maddenin ilişkili olduğu bu iki yetenek arasında telafi edici bir ilişki söz konusudur. Çok boyutlu telafisel modeller, çok boyutluluk için genişletilmiş lojistik modeller ve çok boyutluluk için genişletilmiş normal ogive modeller olmak üzere iki grupta ele alınmaktadır. Çok boyutluluk için genişletilmiş lojistik modeller, (1) çok boyutluluk için genişletilmiş Rasch modeli, (2) çok boyutluluk için genişletilmiş iki parametrelili lojistik model ve (3) çok boyutluluk için genişletilmiş üç parametrelili lojistik model olmak üzere üç grupta ele alınmaktadır. Çok boyutlu lojistik modeller arasında var olan ilişki tek boyutlu MTK'da var olan ilişki ile benzer yapıdadır. Şöyle ki, çok boyutluluk için genişletilmiş Rasch modelinde a parametrelerinin 1 olduğu kabul edilir ve sadece güçlük parametresinin farklılaşması göz önünde bulundurulur. Şans parametresinin olduğu durumda çok boyutluluk için genişletilmiş üç parametrelili lojistik modelin kullanımı tercih edilirken, şans parametresinin sıfır ya da ihmal edilebilir olduğu ve ayırt edicilik parametrelerinin maddeler boyunca farklılaştığı durumda çok boyutluluk için genişletilmiş iki parametrelili lojistik model kullanılır (Çakıcı Eser, 2015; Embretson ve Reise, 2000; Reckase, 2009).

Bu araştırmada kullanılan, çok boyutlu ve 2 parametrelili telafisel modele ilişkin önerilen lojistik olasılık fonksiyonu 4 numaralı eşitlikte verilmiştir:

$$P(U_{ij} = 1 | \theta_j, \alpha_i, d_i) = \frac{e^{\alpha_i \theta_j' + d_i}}{1 + e^{\alpha_i \theta_j' + d_i}} \quad (4)$$

U_{ij} = j bireyinin i maddesine verdiği cevap (1 veya 0)

θ_j = j bireyinin yetenek düzeyi

α_i = i maddesinin ayırtedicilik parametresi

d_i = i maddesinin güçlük parametresi

Çok boyutlu 2 parametrelili lojistik telafisel modelde ayırt edicilik parametresi her boyut için ayrı ayrı verilir. Güçlüğe ilişkin ise sadece bir parametre yer almaktadır.

Çok boyutlu MTK modelinden bir diğeri olan telafisel olmayan (non-compensatory) modelde ise bir maddenin ilişkili olduğu yetenekler arasında telafisellik özelliği söz konusu değildir. Bir diğeri ifadeyle, çok boyutlu yapılarda bir yetenek düzeyindeki yetkinliğin veya becerinin diğeri boyut veya boyutlardaki eksik olan yetkinliği veya beceriyi tamamlamaması durumunda kullanılabilecek bir modeldir.

Telafisel olmayan modellerde ayırt edicilik, güçlük ve yetenek parametresi her boyut için ayrı ayrıdır. Şans parametresi ise telafisel modeldeki gibi her madde için bir tanedir. Bu araştırmada telafisel olmayan model kullanılmayıp çok boyutlu 2 parametrelili lojistik telafisel modelden yararlanıldığı için telafisel olmayan modellere ilişkin sadece kavramsal çerçeve ele alınmıştır.

Literatürde çok sayıda çalışma testlerin boyutluluğunun belirlenmesinin önemini ve gereğini vurgulamaktadır (Finch ve Monahan, 2008; Hattie ve ark.,1996; Ledesma ve Valero-Mora, 2007; Reichenberg, 2013; Özbek, 2012; Sünbül ve Seo, 2012; Svetina, 2011; Svetina ve Levy, 2014; Touron, Lizasoain ve Joarasti, 2012; Yen, 2007). Önerilen boyutluluk belirleme yöntemlerine bakıldığında hiçbirinin tek başına tekboyutluluğa karar vermede yeterli olduğu söylenemez. Bu durum da yapı geçerliği sorununu beraberinde getirmektedir. Bu nedenle bir veri setinin boyutluluğuna birden fazla yöntemle karar vermenin çok daha sağlıklı olacağı düşünülmektedir.

1.1.5. Boyutluluk Belirleme Yöntemleri

Literatürdeki çalışmalar incelendiğinde boyutluluk belirleme yöntemleri genelde parametrik ve parametrik olmayan yöntemler şeklinde ele alındığı görülmektedir (Abswoude, Ark ve Sijtsma, 2004; Mroch ve Bolt, 2006; Özbek, 2012; Reinchenberg, 2013; Svetina, 2011; Svetina ve Levy, 2014). Küçük örneklem, madde sayısının azlığı, boyutlar arasındaki yüksek korelasyon derecesi gibi koşullarda parametrik olmayan yöntemlerin daha iyi sonuçlar vermesi gibi gerekçelerle, parametrik yöntemlerin yanı sıra parametrik olmayan yöntemlerin ve karşılaştırma koşullarının da araştırılıp kullanılması uygun bulunmuştur. Mroch ve Bolt (2006), yaptıkları çalışmada parametrik ve parametrik olmayan boyutluluk belirleme yöntemlerinin karşılaştırılması çalışmalarının alanda yetersiz olduğunu ifade etmiştir ve parametrik olmayan

yöntemlerin model uyumsuzluğu şartları altında parametrik modellere göre avantaj sağlayıp sağlamayacağını bilmesi gerektiğini vurgulamıştır.

Parametrik ve parametrik olmayan yöntemler ile kullanılan bilgisayar yazılımları literatürde çeşitli açılardan gruplandırılmıştır. İkili puanlanan maddelerden oluşan bir testin boyutluluk yapısını belirlemek için kullanılan yöntemler ve bu yöntemlere dayalı analizlerde kullanılan bilgisayar programları Çizelge 2’de sunulmuştur:

Çizelge 2

İkili puanlanan maddelerden oluşan testlerin boyutluluğunu belirlemede kullanılan parametrik ve parametrik olmayan yöntemler ve bilgisayar yazılımları

Parametrik Yöntemler	Bilgisayar Yazılımı
Açımlayıcı Faktör Analizi (Knol ve Berger, 1991)	Mplus, SAS, STATISTICA
Doğrulamalı Faktör Analizi (B.Muthen, 1993)	Mplus
Doğrusal Olmayan Faktör Analizi (McDonald, 1962, 1967)	NOHARM
χ^2 Testi (Gessaroli ve De Champlain, 1996)	CHIDIM
Faktör Analizi (Bock ve Muraki, 1988)	TESTFACT
Parametrik Olmayan Yöntemler	
Hiyerarşik Kümeleme Analizi (Roussos, 1995)	HCA/CCPROX
Boyutluluk Testi (Nandakumar ve Stout, 1993; Stout, 1987)	DIMTEST
Boyutluluk DETECT indeksi (Kim, Zhang ve Stout, 1995)	DETECT
(Tate, 2003, s162)	

Çizelge 2’de yer alan ikili puanlanan maddelerden oluşan testlerin boyutluluğunu belirlemede kullanılan yöntemler ve kim tarafından geliştirildiği ve yöntemlerin kullanımı için geliştirilen bilgisayar yazılım programları ilk geliştirilen yöntemlerden birinin doğrusal olmayan faktör analizi olduğu görülürken bunu takiben faktör analizi ve boyutluluk testi geldiği görülmektedir. Tate (2003)’in bu genel ayrımının yanı sıra Hattie (1985) tekboyutluluğun incelenmesine ilişkin yöntemleri,

güvenirlilik katsayısına bağlı indeksler, faktör analizine dayalı indeksler ve örtük özellik modeline göre gruplandırmıştır.

1.1.5.1. Parametrik Olmayan Yöntemler (Non-parametric Methods)

1.1.5.1.1. Boyutluluk Testi (DIMTEST Süreci)

Boyutluluk Testi (DIMTEST), bir testin tekboyutlu olup olmadığını inceleyen bir istatistiktir. Stout (1987) tarafından geliştirilen bu yöntem boyutluluğa ilişkin H_0 hipotezini test eder.

H_0 : Test tekboyutludur ($d=1$)

H_1 : Test çok boyutludur ($d>1$)

DIMTEST T istatistiği şeklinde de ifade edilen bu yöntem, literatürde tekboyutluluğu belirlemede yaygın olarak kullanılan bir doğrulama sürecidir (Örn, Abswoude, Ark ve Sijtsma, 2004; Mroch ve Bolt, 2006; Özbek, 2012; Özer ve Özkan; 2014; Reichenberg, 2013; Svetina, 2011; Svetina ve Levy, 2014; Tate, 2003). DIMTEST T istatistiği sürecinin gücü, tekboyutluluk alanının keşfedilme gücüne dayanır ve DIMTEST T istatistiği, tekboyutlu ve çok boyutlu testleri ayırmada başarılı kabul edilir ve küçük ikincil özellikleri göstermede bile güçlü bulunur (Nandakumar, 1993; Svetina, 2011).

Bazı çalışmalarda, DIMTEST T istatistiğinin büyük ölçekli örneklemelerde ve geniş madde havuzlarında iyi çalışan bir test etme süreci olduğu ifade edilmektedir (Froelich ve Habing, 2008; Reichenberg, 2013).

DIMTEST T istatistiği uygulamasında veri seti Değerlendirme Alt testleri (Assessment Subtest) (AT1) ve AT2 ile bölümlenme alt testi (Partitioning Subtest) (PT) olmak üzere üç alt sete ayrılmaktadır. Boyutluluk açısından homojen olması gereken madde setlerinden biri AT olarak seçilir. Madde seçim sürecinde madde içeriği, uzman görüşü, faktör analizi, kümeleme analizi gibi analizler temel alınır (Svetina, 2011). Kalan madde kümesi ise PT olarak gruplandırılır. Bu maddeler, testin uygulandığı kişilerin puan gruplarına ayrılması için kullanılır. AT2 alt testi, güçlük açısından AT1 alt testinde yer alan maddelerin güçlüğüne benzer olan maddelerden oluşturulmaktadır. AT2 testinin oluşturulmasının amacı testin güçlüğünün ayrı bir faktör olarak ortaya çıkmasını engellemek diğer bir ifadeyle yanlışlığı önlemektir (Froelich ve Stout, 2003; Nandakumar ve Stout, 1993). DIMTEST T istatistiğinin genel işleyiş süreci kovaryans

matrislerinin eşitliğine dayanmaktadır (Froelich ve Stout, 2003). AT1 ile PT alt testlerinden alınan her bir madde çiftinden elde edilen puanların kovaryanslarının sıfır veya sıfıra yakın olması, testin tekboyutlu olduğu anlamına gelmektedir. Buna göre T_L , AT1 alt testinin varyansına T_B , AT2 alt testinin varyansına dayanarak hesaplanan iki istatistik olmak üzere;

$$T = (T_L - T_B)/(\sqrt{2}) \quad (5)$$

üretilen DIMTEST T istatistiği sıfıra eşit veya yakınsa yokluk hipotezi kabul edilir. Diğer bir ifadeyle testin tekboyutlu olduğu istatistiksel olarak belirlenmiş olur.

DIMTEST T istatistiği 1-0 puanlanan veri setlerinde temelde tekboyutluluğun değerlendirilmesinde bir çerçeve iken yaklaşık basit yapılarla çok boyutlu doğrulayıcı modellerde de boyutluluk değerlendirmelerini sağlamak için kullanılmaktadır (Svetina, 2011).

1.1.5.1.2. DETECT İndeksi (*Dimensionality Evaluation to Enumerate Contributing Traits*)

Parametrik olmayan boyutluluk belirleme yöntemlerinden bir diğeri de DETECT indeksidir. DETECT indeksi, ayrı kümelerdeki test maddelerini sınıflandırmak ve bir testin boyutluluğunun sayısını belirlemek için bir algoritma kullanır (Kim, 1994; Zhang ve Stout, 1999).

DETECT indeksi süreci, test boyutluluk yapısının belirlenmesinde koşullu kovaryanslardan (conditional covariance), yani kalan puanlara bağlı madde çiftlerinin kovaryansından yararlanır. DETECT indeksi, aynı gizil boyutları ölçen maddelerin pozitif koşullu kovaryansa sahip iken farklı gizil boyutları ölçen maddelerin negatif koşullu kovaryansa sahip olduğu prensibine dayalıdır (Zhang ve Stout, 1999). Bu nedenle test yapısının tespit edilmesinde DETECT indeksi 6 numaralı eşitlikte verilen algoritmayı kullanarak indeksi maksimize eden maddelerin ayrımını tespit etmeye çalışır:

$$D(P) = \frac{1}{n(n-1)/2} \sum_{i < j} (-1)(CCov_{x_{ij}} - CCov) \quad (6)$$

Yukarıdaki bağıntıda P , maddelerin belli bir kümelenme şeklini, n toplam madde sayısını, $(CCovx_{ij})$ ise i ve j maddeleri arasındaki şartlı kovaryansı gösterir.

Aslına bakılırsa DETECT indeksi, aynı kümede yer alan maddeleri pozitif koşullu kovaryansa, farklı kümelerdeki maddeleri negatif kovaryansa sahip olacak şekilde ayırmaya çalışır. Burada koşullu kovaryansların büyüklüğünü ağırlık olarak işler.

1.1.5.2. Parametrik Yöntemler (Parametric Methods)

Boyutluluk değerlendirilme sürecinde parametrik yöntemler, faktör analizi veya çok boyutlu MTK'yı temel alarak şekillenmektedir.

1.1.5.2.1. Açımlayıcı Faktör Analizi

Faktör analizi, açımlayıcı boyutluluk analizi ailesinden gelmektedir. Faktör analizi, sürekli veriler üzerinden geliştirilen bir faktör analitik tekniktir. Açımlayıcı faktör analizinde araştırmacı faktör çıkartma yaparak en uygun faktör sayısını belirlemektedir. Faktör analizi, genellikle madde yanıtlarının tek boyutlu MTK modeli ile kalibrasyon (calibration) yapmada yeterli olup olmadığını araştırmak için kullanılır. Faktör analiziyle iki kategorili veriler için tek boyutluluğun test edilmesinde tetrakorik korelasyon matrisi üzerinden işlem yapılmaktadır.

Açımlayıcı Faktör Analizi yöntemi, tekboyutluluğun değerlendirilmesinde yoğun olarak kullanılmasına rağmen örtük yapıyı doğrudan sınıma açısından elverişli olmadığı ifade edilmektedir (Cook, Kallen ve Amtmann, 2009). Doğrulamalı Faktör Analizi (DFA) hipotez test etmeye izin verdiğinden, Açımlayıcı Faktör Analizine göre avantajlı konumda olduğu ifade edilmektedir.

Faktör analitik tekniklerle MTK'da tekboyutluluğun sınanmasında çeşitli sorunlarla karşılaşabileceği vurgulanmaktadır. Bunlardan biri faktör analizinin çok değişkenli normallik gibi belli varsayımlarda bulunmasıdır. Bazı alanlarda bazı veriler normallik varsayımını karşılamayabilir. Faktör analitik yaklaşımların normallik ihlallerine karşı dirençli olduğunu gösteren araştırmalar olsa da başka çalışmalar potansiyel sorunlara işaret eder (Cook, Kallen ve Amtmann, 2009). Mplus ve

TESTFACT programları sürekli veriler için faktör analitik teknikleri kullanarak özdeğerlerden yola çıkarak boyut sayısını belirlemeye çalışmaktadır.

1.1.5.2.2. Doğrusal Olmayan Faktör Analizi: Normal Ogive Harmonic Analysis Robust Method (NOHARM)

Doğrusal Olmayan Faktör Analizi (NOHARM), McDonald (1962, 1967, 1981, 1997) tarafından geliştirip, Fraser ve McDonald (1988) tarafından programlanan açımlayıcı ya da doğrulayıcı analizde kullanılabilen doğrusal olmayan faktör analitik yöntemdir (Svetina, 2011; Tate, 2003). McDonald (1967, 1982), MTK modellerine özel bir doğrusal olmayan faktör analizi durumu gibi ele alarak doğrusal olmayan bir faktör analitik yöntem tanıtmıştır (Seo ve Sünbül, 2012). McDonald, doğrusal olmayan fonksiyonu normal ogive olarak tanımlamıştır. Normal ogive modelin i maddesine doğru yanıt verme olasılığı şöyle tanımlanır (McDonald, 1999):

$$P[U_i = 1 | (\Theta = \theta_1, \theta_2, \dots, \theta_m)] = c_i + (1 - c_i)N(\beta_{i0} + \beta_{i1}\theta_1 + \beta_{i2}\theta_2 + \dots + \beta_{im}\theta_m) \quad (7)$$

Burada $N(.)$ normal dağılım fonksiyonu, Θ m faktörden oluşan aday yeteneği, β_{i0} madde güçlük parametresi, β_{im} m faktörünün madde ayıricılığı parametresi, c_i ise tahmin parametresidir (Seo ve Sünbül, 2012). Fraser tarafından programlanan Doğrusal Olmayan Faktör Analizi (NOHARM) yöntemi hem çok boyutlu MTK hem de faktör model formülasyonları için madde parametreleri kestirimleri sağlar (Tate, 2003). Doğrusal Olmayan Faktör Analizi (NOHARM), açımlayıcı faktör analizinden farklı döndürme tekniklerine (oblique, ortogonal gibi) izin vermektedir (McDonald, 2000). Bu yöntem ve ilgili program şans başarısını (lower asymptotes) kestiremez ama kullanıcının bu değerleri girdi olarak kullanmasına izin verir. Doğrusal Olmayan Faktör Analizi sonuçları, kovaryans artıklarını ve ortalama karekökü vererek uyum eksikliğini özetler. Orijinal halinde olduğu gibi, model uyumuna yönelik bir formal istatistik üretmez (Svetina, 2011).

Boyutluluğun değerlendirmesinde Doğrusal Olmayan Faktör Analizine (NOHARM) dayalı sonuçlar, sonuçların kolay yorumlanabilmesi ve yöntemin madde düzeyinde test yapısının doğrulayıcı analizi için uygun bir işlem olması açısından önemlidir. Ancak birtakım sınırlılıkları bulunmaktadır. Birinci sınırlılık, Doğrusal Olmayan Faktör Analizinde (NOHARM) kullanılan ağırlıklandırılmamış en küçük

kareler kestirim yönteminden dolayı standart hata verilmemesidir. Buna bağlı olarak onunla ilgili işlemi test eden hipotez de yoktur. İkincisi, bir dizi kovaryansın temelini oluşturan boyut sayısını değerlendirecek objektif kriter yokluğudur (Gessaroli & De Champlain, 1996; Jasper, 2010; akt: Sünbül, ve Seo, 2012).

Diğer faktör analizi yöntemlerinde olduğu gibi Doğrusal Olmayan Faktör Analizi de belli bir faktör modeli veya çözümü için çeşitli uyum indeksleri üretir:

1.1.5.2.2.1. Tanaka Uyum İyiliği İndeksi (TIGF)

Doğrusal Olmayan Faktör Analizi için kullanılan NOHARM programının önerdiği TIGF, yapısal eşitlik modellemesi bağlamında yer alan GFI'ye benzer bir indekstir. Bu indeks, verilerin boyutluluğunu ölçmede de kullanılabilir (Tanaka, 1993; akt: Seo ve Sünbül, 2012). TIGFI şu şekilde tanımlanır:

$$TIGF = 1 - \left(\frac{Tr(R^2)}{Tr(S^2)} \right) \quad (8)$$

Burada R , artık kovaryans matrisini; S ise örneklem kovaryans matrisini temsil eder. Genel olarak $TIGF \geq .95$ iyi uyum indeksi için bir gösterge olarak düşünülür (Hu ve Bentler, 1999).

1.1.5.2.2.2. Artıkların Ortalama Kareler Karekökü (RMSR)

Doğrusal Olmayan Faktör Analizi için kullanılan NOHARM programı gözlenen ve beklenen oranlarla RMSR arasındaki farkların artık matrisini verir (Stone ve Yeh, 2006). RMSR aşağıdaki Formül 9'dan elde edilir:

$$RMSR = \sqrt{2 \sum_{i=2}^p \sum_{j=1}^{i-1} \frac{(r_{ij})^2}{p(p-1)}} \quad (9)$$

Burada p madde sayısını; r_{ij} , i ve j ($i, j = 1, \dots, p$) maddeleri arasındaki artık kovaryansını temsil eder.

1.1.5.2.2.3. Yaklaşık Ki-Kare ($\chi^2_{G/D}$)

Doğrusal Olmayan Faktör Analizi için kullanılan NOHARM programına dayalı belli bir boyutluluk çözümünün uyumunun iyiliğini test edecek formal bir örnek de Gessaroli ve De Champlain (1996) tarafından ki kare istatistiği olarak tanıtılmıştır. “Faktör analizinden elde edilen kalan korelasyon matrisindeki köşegen dışı öğeler sıfıra eşittir.” şeklindeki sıfır hipotezinin sınanmasına dayalı bir istatistiktir (Finch ve Habing, 2005). Eğer sıfır hipotezi reddedilmezse uyumlu modelin maddeler arasında gözlemlenen korelasyonları yeterince tahmin ettiği sonucuna varılabilir (Finch ve Habing, 2005). Yaklaşık ki-kare istatistiği Formül 10’daki şekilde hesaplanabilir:

$$\chi^2_{G/D} = (N - 3) \sum_{i=2}^p \sum_{j=1}^{i-1} Z_{ij}^{2(r)} \quad (10)$$

Formül 11’e göre N , kişi sayısını, J toplam madde sayısını temsil eder, j ve j' ise özel madde çiftlerini tanımlayacak maddeleri indekslemeye yardımcı olur. $Z_{ij}^{2(r)}$, i ve j , ($i, j = 1, \dots, p$) maddeleri arasındaki artık korelasyona karşılık gelen Fisher Z ’nin karesini ifade etmektedir. Ortaya çıkan istatistik, $sd=0.5J(J-1)-t$ serbestlik derecesine sahip referans ki-kare dağılımı ile kıyaslanır. Burada t , model uyumunda kestirimi yapılan serbest parametre sayısını gösterir; açıklayıcı modellerde $t=(1+m) \times J$ olup m , boyut sayısıdır (Finch ve Habing, 2007).

Bu araştırmada DIMTEST T istatistiğinin tercih edilmesinin en önemli nedeni büyük örneklemelerde ve geniş madde havuzlarında iyi çalışan bir test etme süreci olduğu ve küçük ikincil özellikleri bile göstermede etkili bulunduğu içindir. Doğrusal Olmayan Faktör Analizi’nin tercih edilmesinin nedeni ise sonuçların kolay yorumlanabilmesi, küçük örneklemelerde iyi çalışması ve faktör analitik yaklaşımları temel almasından kaynaklamaktadır. Bütün bunların yanı sıra tüm yöntemlerin ücretsiz erişilebilir olması tercih edilme nedenini güçlendirmiştir.

1.2. İlgili Araştırmalar

İlgili alanyazın incelendiğinde, araştırmacıların çeşitli boyutluluk belirleme yöntemlerinin performansını incelediği görülmektedir. Boyutluluk belirleme

yöntemlerinin belli koşullar altında karşılaştırma yapan çalışmaların yanı sıra boyutluluk belirleme yöntemlerinin tek boyutluluğa ya da çokboyutluluğa karar verme sürecinde kullanıldığı çalışmalar da mevcuttur. Bu bölümde yöntemlerin belli koşullar altında karşılaştırıldığı çalışmaların yanı sıra ülkemizde yapılan ve bu yöntemlerin çeşitli araştırmalarda kullanıldığı bazı ulusal çalışmalara yer verilmiştir.

1.2.1. Boyutluluk Yöntemlerin Karşılaştırıldığı Çalışmalar

Gessaroli ve De Champlain (1996) tarafından yapılan araştırmada, doğrusal olmayan faktör analizi (NOHARM) indekslerinden yaklaşık ki kare istatistiği ($\chi^2_{G/D}$) ile DIMTEST T istatistiği karşılaştırılmıştır. İki parametrelili lojistik (2PL) modeller için 1. tip hata oranının her iki yöntem için de neredeyse eşdeğer olduğu tespit edilmiştir. Test uzunluğu ile ilgili olarak ise test maddeleri az sayıda (15) ya da orta düzeyde (30) iken doğrusal olmayan faktör analizi sonucu 1. tip hata oranının DIMTEST T istatistiğine göre çok daha düşük olduğu ve madde sayısı büyük iken (45) ise DIMTEST T istatistiği kadar iyi performans gösterdiği ifade edilmiştir.

Zhang ve Stout (1999) yaptığı çalışmada boyut sayısını (2, 3 ve 4), test uzunluğu (20 ve 40) ve örneklem sayısını (400 ve 800) olacak şekilde değişimleyerek verileri üretmiştir. Her bir koşul 100 kere tekrar edilmiştir. Çokboyutlu telafisel model kullanılarak yaklaşık basit yapı gösteren veriler üretilmiştir. Yazarlar örneklem büyüklüğü arttıkça DETECT'in performansının arttığını ortaya koymuşlardır. Örneklem büyüklüğü ve madde sayısı sabit tutulup boyut sayısı arttırıldığında DETECT performansının zayıfladığı görülmüştür. Performanstaki bu zayıflama en çok 4 boyutlu 20 maddelik ve örneklem büyüklüğü 400 olduğu durumda görülmüştür. Çalışma kapsamında bir diğer alt çalışma tek boyutlu durumlarla ilgilidir. Test uzunluğu (20 ve 40 madde), örneklem büyüklüğü (400 ve 800) ve şans parametresinin değeri (.00 ve .20) olacak şekilde koşullar manipüle edilmiştir. DETECT testlerin tüm durumları için tek boyutlu olduğunu doğrulamada başarılı olduğu gözlenmiştir. Yaklaşık basit yapıda ise baskın örtük boyutların belirlenmesi ve testte yer alan çokboyutluluk miktarının kestirilmesinde DETECT'in iyi çalıştığı ortaya konulmuştur.

Alexandra, Abswoude, Ark ve Sijitma (2004) tarafından, parametrik olmayan IRT modellerinde, DETECT, DIMTEST ve HCA/CCPROX yöntemleri karşılaştırılmıştır. Çalışma kapsamında, MTK modeli (2 PL ve 5 PAM), madde seçim

yöntemi (4 yöntem), madde ayırtedicilikleri, her bir boyuttaki madde sayısı, boyutlararası korelasyon (0.0, 0.2, 0.4, 0.6, 0.8, 1.0) değişimlenmiştir. DIMTEST ve DETECT'i temel alan istatistiklerin diğer istatistiklere göre tek boyutluluğu kabul ya da reddetmedeki doğruluğunun dikkat çekici olduğu ifade edilmiştir.

Twu (2005), Stout'un DIMTEST T istatistiği ve McDonald'ın Doğrusal Olmayan Faktör Analizi (NOHARM) olmak üzere ölçme uygulamalarında tek boyutluluk belirleme sürecinde başvuru olan iki yöntemi karşılaştırılmıştır. Çalışma kapsamında, tek boyutlu, iki boyutlu ve yaklaşık tek boyutlu veri setleri olmak üzere üç farklı veri türü simüle edilmiştir. Her veri türü için dört farklı örneklem büyüklüğü (200, 500, 1000, 2000 öğrenci), iki farklı test uzunluğu (25 madde, 50 madde) ve iki şans başarısı düzeyi ($c = 0$, $c \neq 0$) manipüle edilmiştir. Yaklaşık tek boyutlu veri setleri için özellikler arasında iki farklı korelasyon seti kullanılmıştır. Her faktör kombinasyonunda 500 replikasyon gerçekleştirilmiştir. DIMTEST T istatistiğinde AT1 maddelerinin seçiminde üç yöntem (faktör analizi, tesadüfi seçim ve sistematik seçim) kullanılmıştır. Doğrusal Olmayan Faktör Analizi (NOHARM) sonuçlarının yorumlanmasında üç indeks (RMSR, Tanaka Uyum İyiliği İndeksi ve yaklaşık ki kare istatistiği) kullanılmıştır. Araştırma sonucunda, DIMTEST T istatistiğinin tek boyutluluk belirlemede belli bir oranda tutarlılık elde edilmiştir. AT1 maddelerinin seçilme şekli DIMTEST T istatistiği sonuçları üzerinde etkili olduğu ifade edilmiştir. Örneklem büyüklüğü arttıkça veya testlerin uzunluğu arttıkça DIMTEST T istatistiği daha yüksek ret oranları göstermiştir; boyutlararası korelasyonların yüksek olduğu durumlarda DIMTEST T istatistiğinin ret oranları düşük çıkmıştır. Doğrusal Olmayan Faktör Analizi (NOHARM) ise tek boyutluluktan ayrılmayı doğru tespit etmiş ve çok boyutluluğa ilişkin karar vermede DIMTEST T istatistiğinden daha duyarlı sonuçlar elde edilmiştir. Örneklem büyüklüğü, test uzunluğu, şans başarısı ve boyutlararası arasındaki korelasyonun NOHARM performansına etkisi manidar düzeyde çıkmamıştır.

Roussos ve Özbek (2006) tarafından yapılan çalışmada, çok büyük örneklem (120.000) kullanılarak DETECT'te bulunan istatistiksel yanlışlık miktarı değerlendirilmiştir. Çalışma kapsamında, boyut sayısı (1, 2, 3), boyutlar arası korelasyonlar (.50, .70), her boyuta ait test uzunluğu (tek boyutluda 5 ile 40 arasında, çokboyutluda 20 ile 40 madde arasında) ve veri yapısı (basit ya da yaklaşık basit yapı) şeklinde değişimlemeye gidilmiştir. Araştırmacılar DETECT yönteminin özellikle 10 ya

da daha az test uzunluęu olduęu ve tekboyutlu olduęu durumlarda istatistiksel yanlılık sergiledięini söylemektedir . Buna göre arařtırmacılar, 20 maddeden az test uzunluęu için DETECT kullanmamayı önermiřlerdir. Çok boyutlu durumlarda yazarlar büyük örneklemli DETECT indeksinin tüm simülasyon kořulları için oldukça küçük yanlılık gösterdięini belirtmiřlerdir. Ayrıca DETECT'in yüksek doęruluk oranı saęladıęı görölmüřtür. 45 çok boyutlu kořulun sadece 3 tanesi mükemmel doęruluk oranından (1.00) daha azını göstermiřtir. Dięer geri kalanları mükemmel doęruluk oranına sahiptir (1.00). En düşük sınıflandırma oranı 20 maddelik yaklaşık basit yapılı iki boyutlu kořullar için 0.91 doęruluk oranı saęlamıřtır.

Finch ve Habing (2007) yaptıęı arařtırmada, tekboyutluluęun deęerlendirilmesinde, DIMTEST ve NOHARM'ı temel alan istatistiklerin performanslarını karřılařtırmıřlardır. Çalışma kapsamında DIMTEST'in parametrik bootstrap versiyonu ile NOHARM'ın üç uyum indeksini karřılařtırılmıřtır. Bu kapsamda madde sayıları (15, 30, 60), MTK Modeli (2PL ve 3PL), daęılımın çarpıklıęı (-1,5; -0,5; 0,0; 0,5; 1,5) ve örneklem büyüklüęü (1000, 2000) bulunduęu durumlar arařtırmaya dahil edilmiřtir. řans parametresinin olmadıęı durumlarda, DIMTEST ve NOHARM temelli istatistiklerin benzer güç oranları sergiledięi, çok düşük 1. tip hata gösterdięi ifade edilmiřtir. Buna karřın řans parametresi olduęu durumlarda NOHARM istatistiklerinin beklenmedik bir řekilde yüksek çıktıęı ifade edilmiřtir.

Yu, Popp, DiGangi ve Pennell (2007), tekboyutluluęu deęerlendirmede Rash Modelleme, Paralel Analiz ve TETRAD yöntemlerini karřılařtıran bir çalışma yapmıřtır. Faktör analitik yöntemlerin boyutluluęu belirlemede yetersiz olduęunu, paralel analizin çok az bilinmesine ve istatistik paket programında yer almamasına raęmen boyutluluęu belirlemede en bařarılı yöntemlerden biri olduęu ifade edilmektedir.

Finch ve Monahan (2008), Modifiye Edilmiř Paralel Analiz (Modified Paralel Analysis) ve DIMTEST analizini karřılařtırmıřtır. Çalışma kapsamında, 1. Tip hata ve testin gücü karřılařtırılmıřtır. Testin uzunluęu 15-30-60, örneklem büyüklüęü 500-1000-2000 ve boyutlar arası korelasyon 0- 0,3- 0,5- 0,7- 0,8 ve 0,9 olarak belirlenmiřtir. Modeller 2 PL ve 3 PL olarak belirlenmiřtir. Testin gücü belirlenirken iki boyutlu MTK, 1.tip hata oranını belirlenirken tek boyutlu MTK modelinden yararlanılmıřtır DIMTEST T istatistięini elde etmek için DIMTEST kullanılırken BMPAT istatistięini

elde etmek için ise TESTFACT kullanılmıştır. Bootstrap testinin, 1. tip hatayı kontrol etmede DIMTEST'e göre daha etkili olduğu gözlenmiştir. DIMTEST örneklem büyüklüğünden önemli derecede etkilendiği görülmüştür. 1. Tip hata oranı DIMTEST'te küçük örneklemelerde artarken, örneklem sayısı arttıkça azaldığı ifade edilmiştir. BMPAT'nin ise ne örneklem sayısından ne de testin uzunluğundan etkilendiği görülmüştür. Örtük özellikler arasında korelasyon miktarı arttıkça BMPAT'de sıkıntı olduğu ifade edilmiştir. ($r=0,7 - 0,8$ arası). Birey sayısı ve testin uzunluğu arttıkça testin gücünün arttığı görülmüştür. Bunun yanı sıra bu çalışmadan çıkarılabilecek en büyük sonuç, birçok gerçek durumda MPA bootstrap testi, testlerde çokboyutluluğu belirlemek için “Marjinal Maximum Likelihood Faktör Analizi” ni kullanarak faydalı bir alternatif sunması olarak ifade edilmektedir. Bu çalışmada maddeler ve örtük özellikler “basit yapı” olarak düşünülmüştür ve daha sonraki araştırmalar için karmaşık yapılarla ilgili koşulların da sınanması gerektiği vurgulanmıştır. Bu tarz çalışmaların başarı testi dışındaki alanlarda (yaşam kalitesi ölçeği, kişilik testi gibi) küçük ölçekli test durumları için de uygulanması önerilmiştir. Son olarak da örtük özelliklerin normal dağılmadığı durumlarda da test edilmesi gerektiği vurgulanmıştır.

Svetina (2011), DETECT ve NOHARM boyutluluk belirleme yöntemlerini kullanarak telafisel ve telafisel olmayan çok boyutlu madde tepki modellerindeki boyutluluğun değerlendirilmesinde karmaşık yapıların etkisini araştırmıştır. DETECT'i temel alan yöntemlerin NOHARM'ı temel alan 2 ve 3 boyutlu madde tepki modellerine göre daha iyi çalıştığını ifade etmektedir. DETECT'in özellikle büyük örneklemelerde, % 30 veya daha az karmaşık verilerde ve boyutlar arası korelasyonun 0.60 veya daha az olduğu durumlarda daha doğru oranlar sağladığını ifade edilirken, DETECT temelli istatistiklerin örneklem büyüklüğü azaldıkça, karmaşık yapı arttıkça performansının azaldığı, DETECT'i temel alan yöntemlerin karmaşık yapı, madde sayısı ve boyutluluk arttıkça NOHARM'ı temel alan yöntemlere göre daha tutarlı sonuçlar verdiği ifade edilmiştir. Bunun yanı sıra DETECT'in tutarlılığına rağmen, telafisel olmayan çok boyutlu MTK modellerine uygun verilerde, boyutluluğu belirlemede NOHARM'ı temel alan istatistiklerin DETECT'e göre daha güçlü performans gösterdiği ifade edilmektedir.

Seo ve Sünbül (2012) tarafından, DIMTEST ve NOHARM temelli istatistiklerin performanslarını karşılaştırmak amacıyla örneklem büyüklüğü (1000, 2000, 3000, 5000), boyutlararası korelasyon (0.0, 0.3, 0.8 ve 0.95), DIMTEST AT-PT oranları (5:5, 7:3, 8:2) ve testin uzunluğu (15, 30, 60) değişimlenerek bir çalışma yapılmıştır. DIMTEST'in tek başına kullanıldığında büyük örneklemelerde kısa testler için yüksek oranda I.tip hata verdiği belirtilmiştir. DIMTEST 0.95 korelasyonu için bile çokboyutlu verilerde yüksek ret oranları çıkardığı ifade edilmiştir. NOHARM'ın RMSR veya TIGF uyum indekslerinde ise test verilerinin tekboyutluluğunun doğrulanmasında iyi model uyumu hakkında doğru bir gösterge olmadığı belirtilmiştir.

Svetina ve Levy (2014) tarafından, çok boyutlu MTK modellerinde boyutluluğun değerlendirilmesi süreci ile ilgili genel bir çerçeve sunulmuştur. Bu çerçeve özellikle açıklayıcı ve doğrulayıcı yaklaşımları, parametrik ve parametrik olmayan varsayımları, 1-0, çoklu ve kayıp verili uygulamaları içeren bir süreçtir. Bu kapsamda öneriler Eğitim Sürecindeki Ulusal Değerlendirme Biliminin (NAEP) verileri analiz edilmiştir ve karşılaştırılmalara gidilmiştir.

1.2.2. İlgili Yöntemlerin Kullanıldığı Ulusal Çalışmalar

Özbek Baştuğ (2012) un yaptığı çalışmada, Orta Okullar Arası Seçme Sınavı'nın Sosyal Bilgiler alt testinin boyutları parametrik ve parametrik olmayan çok boyutluluk yöntemleri birlikte kullanılarak tanımlanmıştır. 1999, 2000 ve 2001 Orta Okullar Arası Seçme Sınavının Sosyal Bilgiler Testinin verileri için bu çalışmada özellikle parametrik ve parametrik olmayan çokboyutluluk metotlarından (DIMTEST, ATFIND, DETECT, HCA/CCPROX), doğrusal olmayan faktör analizine dayanan çok boyutluluk metodu (NOHARM) kullanılmıştır. HCA/CCPROX'un bulgusu, DIMTEST ve DETECT analizi sonuçları uyumlu çıktığı ve çok boyutluluğu açıklamada doğru olduğu ifade edilmiştir. Ancak altta yatan gizil çok boyutlu yapıyı bulmak için faydalı bir metot olduğu kanıtlanan NOHARM'ın, bu araştırmada güvenilir sonuçlar vermekte başarısız olduğu ifade edilmiştir.

Özer Özkan ve Acar Güvendir (2014) tarafından, Öğrenci Başarısının Belirlenmesi Sınavının (ÖBBS) dört alt testinin boyutları parametrik olmayan çokboyutluluk yöntemlerinden DIMTEST T istatistiği kullanılarak belirlenmiştir. Hipotez testi sonucu 2002, 2005 ve 2008 yıllarında uygulanan ÖBBS'nin Türkçe,

matematik, fen ve teknoloji, sosyal bilgiler ve İngilizce tüm alt testlerinin çokboyutlu olduğu olduğu bulgusuna ulaşılmıştır.

Kaya ve Kelecioğlu (2016) tarafından yapılan araştırmada, örneklem büyüklüğünün test boyutluluğuna etkisi incelenmiştir. Boyutluluğun belirlenmesinde parametrik ve parametrik olmayan açımlayıcı faktör analizi yöntemlerinden olan DIMTEST, DETECT, NOHARM ve Paralel Analiz'den yararlanılmıştır. Örneklem büyüklüğü 50, 80, 100, 160, 200, 300, 500, 1000, 3000, 5000 şeklinde değişimlenmiştir. DETECT ve NOHARM'ın çok boyutlulukta örneklem sayısının 50 olduğu durumda kabul edilebilir düzeyde iken, 1000 ve üzeri örneklem büyüklüklerinde daha tutarlı sonuçlar ürettiği ifade edilmektedir. Tetrakorik korelasyon matrisini temel alan Paralel Analiz'de ise daha tutarsız sonuçlar elde edildiği ifade edilmiştir.

İlgili alan yazın incelendiğinde, tekboyutluluğu belirlemede kullanılan yöntemlerin karşılaştırılması ile ilgili araştırmalar olduğu görülmektedir. Araştırmalara bakıldığında, çalışmaların genelde standart normal dağılım gösteren veri setlerinde ve büyük örneklemeler üzerinden karşılaştırıldıkları, çarpık dağılım gösteren veriler ve küçük örneklemeler için genelde tercih edilmemektedir. Sosyal bilimlerde kullanılan testlerin dağılımları gözününde bulundurulduğunda ile okul uygulamalarındaki örneklem sayısının azlığı ve tüm bu koşullarda yöntemlerin karşılaştırılması bu çalışmanın gereği olarak görülmüştür.

1.3. Amaç

Bu araştırmanın genel amacı tekboyutlu ve iki boyutlu yapılarda boyutluluk belirlenmesinde kullanılan yöntemlerin 1. tip hata ve güç oranlarının çeşitli koşullar için araştırılmasıdır ve her bir koşulun temel etkisinin yanı sıra koşulların ortak etkisini (koşulların etkileşiminin etkisi) karşılaştırmaktır. Bu genel amaç doğrultusunda aşağıdaki sorulara cevap aranmıştır;

1. 1-0 puanlanan testlerde çeşitli boyutluluk belirleme yöntemlerine göre testin uzunluğunun, dağılımın özelliğinin ve örneklem büyüklüğünün manipule edildiği tek boyutlu verilerden elde edilen *1. tip hata oranları* nasıl değişmektedir?

2. 1-0 puanlanan testlerde çeşitli boyutluluk belirleme yöntemlerine göre testin uzunluğunun, boyutlararası korelasyon derecesinin, dağılımların ve örneklem büyüklüğünün manipule edildiği iki boyutlu verilerden elde edilen *testin güç oranları* nasıl değişmektedir?
3. 1-0 puanlanan testlerde çeşitli boyutluluk belirleme yöntemlerine göre standart normal ve çarpık dağılım gösteren verilerde 1. Tip hata oranı ve testin güç oranları nasıldır?

1.4. Önem

Test ve ölçek geliştirme sürecinde ölçülecek özelliğin boyutluluğunun değerlendirilmesi geliştirilen testin ya da ölçeğin geçerliğine kanıt sağlaması açısından önem taşımaktadır. Boyutluluğun değerlendirilmesi özellikle tekboyutlu MTK uygulamaları için son derece önemlidir. Tekboyutlu MTK uygulamalarında tek boyutluluk sayıtlısının karşılanmaması yerel bağımsızlık sayıtlısını beraberinde zedeleyecektir. Bu durum da test eşitleme hatalarının yanı sıra yanlı parametre kestirimlerine neden olacaktır.

Tekboyutluluk incelemelerinde faktör analizi genellikle tercih edilmesine rağmen birçok çalışmada sadece faktör analizi ile tek boyutluluk incelenmesinin yeterli olmadığı ifade edilmiş olup bu incelemelere başka yöntemler de önerilmektedir (Finch ve Monahan, 2008; Hattie ve ark., 1996; Ledesma ve Valero-Mora, 2007; Özbek, 2012; Reichenberg 2013; Svetina, 2011; Svetina ve Levy, 2014; Touron, Lizasoain ve Joarasti, 2012; Yen, 2007) fakat buna rağmen ulusal ya da uluslararası birçok araştırmada tekboyutluluğun incelenmesinde hala sadece faktör analizi kullanılmaktadır ve yeterli görülmektedir. Oysaki faktör analizi çok değişkenli normal dağılım varsayımını gerektirmektedir ve sosyal bilimlerde çoğu zaman bu sağlanamayabilir.

Ulusal ya da uluslar arası test merkezleri tarafından kullanılan başarı testlerinin sadece faktör analizi kullanılıp ya da herhangi bir yöntem kullanılmadan tekboyutlu kabul edilmesi ve test puanlarının toplam test üzerinden hesaplanıp bireyler hakkında karar verilmesi –test ve ölçek geliştirmenin doğası gereği- verilen kararların geçerliği açısından bir sınırlılık teşkil etmektedir. Tekboyutluluğun ihlalinin varsa çok boyutlu yapının doğru yöntemler ve indekslerle tespit edilmesi yapı geçerliği çalışmaları için mutlaka araştırılması gerektiği düşünülmektedir.

Tekboyutluluk belirleme sürecinde bir diğer önemli görülen nokta ise, ülkemizde veri toplama süreçlerinde yaşanan zorluklar gözönüne alındığında, test geliştiriciler için, örneklem büyüklüğünün tekboyutluluk belirlemedeki etkisinin mutlaka araştırılması gereken bir konu olduğu düşünülmektedir.

İlgili alanyazın incelendiğinde boyutluluk belirleme sürecinde örneklem büyüklüğü, testin uzunluğu, veri türü (1-0 ve çoklu puanlanan), boyutlararası korelasyonun derecesi gibi değişkenlerin manipule edildiği çalışmalar olmakla beraber küçük örneklemelerde ve test uzunluğunun az olduğu koşullarda dağılımın türünün (normal dağılım-çarpık dağılım) manipule edildiği herhangi bir çalışmaya rastlanmamıştır. Bunun yanı sıra manipule edilecek koşulların temel etkilerinin (her bir koşulun etkisi) incelendiği uluslararası düzeyde çalışmalar olduğu görülmüştür ancak koşulların ortak etkilerinin (koşulların etkileşimin etkisi) incelendiği çok az sayıda çalışmaya rastlanmıştır. Tüm bunların yanı sıra üç yöntem ve bu yöntemlere bağlı olan dört indeksin bir arada karşılaştırıldığı herhangi bir çalışmaya rastlanmamıştır. Bu nedenle bu çalışma boyutluluk belirlenmesinde faktör analizi dışında kullanılabilecek çeşitli yöntem ve indekslerin tanınması açısından önemli görülmektedir.

1.5. Sınırlılıklar

Boyutluluk belirleme yöntemleri, kullanılan analiz programları önerideki yöntemlerle sınırlıdır.

Araştırmada çokboyutluluk iki boyutlu yapılarla sınırlıdır.

BÖLÜM II

2. YÖNTEM

Bu bölümde; araştırmanın modeli, veri üretim çalışması ve verilerin analizi açıklanmıştır.

2.1. Araştırmanın Modeli

Bu araştırma, tekboyutluluğun belirlenmesinde kullanılan yöntemlerin 1. Tip hata ve güç oranlarını belirlemeye yönelik ve boyutluluk belirleme yöntemlerini karşılaştırma olanağı da verdiği için kuramsal bir araştırma olduğu söylenebilir.

2.2. Veri Üretim Çalışması

Bu çalışmada araştırma sorularına yanıt vermek için simülasyon verisinden yararlanılmıştır. Simülasyon modelleri gerçekçi parametrelere dayanmalıdır (Davey, Nering ve Thompson, 1997; akt:Göçer Şahin, 2016). Bunun yanı sıra simülasyon çalışmaları gerçek duruma benzediği koşulda anlamlıdır. Bu çalışmada belirtilen tüm koşulların aynı anda gerçek veride karşılanmasının zor olması nedeniyle simülasyon verisinin kullanılması tercih edilmiştir.

Bu çalışmanın verileri Ek 1 ve Ek 2’de yer alan kodlar aracılığı ile SAS programı kullanılarak üretilmiştir. Veriler, güç çalışması için 2 parametrelili lojistik ve telafisel modelde, iki kategorili (1-0) 2 boyutlu yapıya uygun olacak şekilde üretilmiştir. 1. tip hata çalışması için ise 2 parametrelili lojistik modelde, tekboyutlu iki kategorili (1-0) veri üretilmiştir.

Bu araştırma 4 değişken içermektedir. *a) dağılımın özellikleri, b) örneklem büyüklüğü, c) test uzunluğu, d) boyutlar arası korelasyon,*

a) Dağılımın özellikleri: Zeka, tutum gibi psikolojik özellikler evrende simetrik normal dağılım göstermektedir (Glenberg, 1996; akt: Slocum-Gori ve Zumbo,

2010). Buna karşın sosyal bilimlerde ve davranış bilimlerindeki çalışmaların dağılımı sıklıkla çarpık dağılım özelliği göstermektedir (Aron ve Aron, 2002; akt: Slocum-Gori ve Zumbo, 2010). Bu nedenle bu çalışmada kullanılan yöntemler hem standart normal dağılım hem de çarpık dağılıma göre incelenmiştir. Bireylerin yetenek parametreleri $N(0,1)$ standart normal dağılımından ve çarpık dağılım için Fleisman'ın (1978) çalışmasında önerilen çarpıklık basıklık değerleri (1.75, 3.75) temel alınarak üretilmiştir.

- b) Örneklem büyüklüğü:** Son yıllarda DETECT IDN İndeksi, Boyutluluk DIMTEST T istatistiği, Faktör Analizi, Hiyerarşik Kümeleme Analizi gibi boyutluluk belirleme yöntemlerinin performansını araştıran çalışmalarda örneklem büyüklüğü 500, 1000 ve 2000 dahil olacak şekilde belirlenmiştir (Finch ve Habing, 2005; Gessaroli ve De Champlain, 1996; Svetina, 2011). Bu çalışmada ise 200, 300, 500, 1000, 2000 ve 3000 şeklinde örneklem kullanılmıştır.
- c) Testin uzunluğu:** Testin uzunluğunun yöntemlere etkisini inceleyebilmek için özellikle ülkemizde kullanılan başarı testlerinde yer alan alt testlerin genel olarak uzunluğu da dikkate alınarak madde sayısı 10, 20 ve 30 olarak belirlenmiştir.
- d) Boyutlar arası korelasyon:** Testin gücüne ilişkin çeşitli kararlar verilebilmesi için boyutlar arası korelasyonu göz önünde bulundurmak gerekir. Bir test çok boyutlu iken çok boyutlu kararını vermek testin gücünün göstergesidir. Boyutlar arası korelasyon arttıkça yapının tekboyutluluğa yaklaştığı söylenebilir. Bu nedenle literatürde yer alan boyutluluk çalışmalarında (Svetina, 2011; Svetina ve Levy, 2014) –karşılaştırma yapma olanağı açısından- boyut sayısı ve boyutlar arası korelasyonun dereceleri de göz önüne alınarak bu çalışmada iki boyutlu yapılar oluşturulup boyutlar arası korelasyon dereceleri $r=0.25$, $r=0.50$, $r=0.75$, $r=0.90$ olarak belirlenmiştir.

Veri üretiminde kullanılan değişkenler, koşul sayısı ve koşul değerleri Çizelge 3'de sunulmuştur:

Çizelge 3.

Veri üretiminde kullanılan değişken ve koşulları

Yapılan Çalışma	Değişkenler	Koşul Sayısı	Koşul Değerleri
I. Tip hata çalışması	Dağılımın Özellikleri	2	Normal, Çarpık
	Örneklem Büyüklüğü	6	200, 300, 500, 1000, 2000 ve 3000
	Testin Uzunluğu	3	10, 20, 30
Güç Çalışması	Dağılımın Özellikleri	2	Normal, Çarpık
	Örneklem Büyüklüğü	6	200, 300, 500, 1000, 2000 ve 3000
	Testin Uzunluğu	3	10, 20, 30
	Boyutlararası korelasyon	4	0.25, 0.50, 0.75, 0.90
Replikasyon Sayısı		100	

Çizelge 3 incelendiğinde, manipüle edilen değişkenler göz önünde bulundurularak 1. tip hata çalışması için; $2*6*3= 36$ koşul ve güç çalışması için ise; $2*6*3*4=144$ koşul üretilmiştir ve her koşul için 100 replikasyon gerçekleştirilmiştir. 1. tip hata çalışması için araştırma deseni Çizelge 4’de yer almaktadır:

Çizelge 4.

1. tip hata için araştırma deseni

Örneklem Büyüklüğü	Madde Sayısı	Dağılımın Türü	
		normal	çarpık
200	10	U1	U2
	20	U3	U4
	30	U5	U6
300	10	U7	U8
	20	U9	U10
	30	U11	U12
500	10	U13	U14
	20	U15	U16
	30	U17	U18
1000	10	U19	U20
	20	U21	U22
	30	U23	U24
2000	10	U25	U26
	20	U27	U28
	30	U29	U30
3000	10	U31	U32
	20	U33	U34
	30	U35	U36

Çizelge 4 incelendiğinde, 1. Tip hata için, değişkenlerin koşul değerlerinin manipule edildiği 36 hücre oluşturulmuştur. Güç çalışması için araştırma deseni Çizelge 5’de yer almaktadır:

Çizelge 5.

Güç çalışması için araştırma deseni

		Boyutlar Arası Korelasyon							
		0.25		0.50		0.75		0.90	
Örneklem Büyüklüğü	Madde Sayısı	normal	çarpık	normal	çarpık	normal	çarpık	normal	çarpık
200	10	D1	D2	D3	D4	D5	D6	D7	D8
	20	D9	D10	D11	D12	D13	D14	D15	D16
	30	D17	D18	D19	D20	D21	D22	D23	D24
300	10	D25	D26	D27	D28	D29	D30	D31	D32
	20	D33	D34	D35	D36	D37	D38	D39	D40
	30	D41	D42	D43	D44	D45	D46	D47	D48
500	10	D49	D50	D51	D52	D53	D54	D55	D56
	20	D57	D58	D59	D60	D61	D62	D63	D64
	30	D65	D66	D67	D68	D69	D70	D71	D72
1000	10	D73	D74	D75	D76	D77	D78	D79	D80
	20	D81	D82	D83	D84	D85	D86	D87	D88
	30	D89	D90	D91	D92	D93	D94	D95	D96
2000	10	D97	D98	D99	D100	D101	D102	D103	D104
	20	D105	D106	D107	D108	D109	D110	D111	D112
	30	D113	D114	D115	D116	D117	D118	D119	D120
3000	10	D121	D122	D123	D124	D125	D126	D127	D128
	20	D129	D130	D131	D132	D133	D134	D135	D136
	30	D137	D138	D139	D140	D141	D142	D143	D144

Çizelge 5 incelendiğinde, güç çalışması için, değişkenlerin koşul değerlerinin manipule edildiği 144 hücre oluşturulmuştur. Verilerin üretiminden önce maddelerin ayırt edicilik parametreleri araştırmanın deseni göz önünde bulundurularak oluşturulmuştur. Testin çokboyutluluğu ayırt edicilik parametrelerine göre belirlenmiştir. Buna göre her iki boyuta yük veren bir maddenin iki ayırt edicilik katsayısı bulunması gereklidir. Eğer madde her iki boyuta baskın olarak yük veriyorsa bu madde karmaşık; bir boyuta baskın diğerine az bir miktarda yük veriyorsa yaklaşık basit, bir boyuta baskın diğer boyuta hiç yük vermiyorsa basit madde olarak nitelendirilmektedir. Bu çalışmada örneğin 10 maddelik bir testin ilk beş maddesi baskın olarak birinci boyuta ve az bir miktarda ikinci boyuta; diğer beş maddesi ise baskın olarak ikinci boyuta az bir oranda ise birinci boyuta yük verecek şekilde

düzenlenmiştir. Böylelikle farklı iki boyuta baskın olarak yük veren çokboyutlu bir test oluşturulmuştur. Madde parametreleri üretilirken Ackerman (1994) tarafından geliştirilen ITEMGENv2 programından yararlanılmıştır. Bu programda yalnızca, dosya adı, testin uzunluğu, maddelerin açıları, arakesit parametresinin ranjı, MDISC parametresinin ranjına ilişkin bilgiler girilerek parametreler üretilmektedir. Buna göre birinci boyuta yük veren maddelerin x eksenine yaptığı açı 5° - 20° arasında, ikinci boyuta yük veren maddelerin açıları ise 70° ile 88° arasında değişmektedir (Ackerman, Gierl ve Walker, 2003).

MDISC, çok boyutlu MTK'ya ait ayırt edicilik parametresi olup, tekboyutlu MTK'deki madde ayırtediciliğine karşılık gelir. Burada birden fazla boyut yer aldığı için her boyut için bir ayırt edicilik söz konusudur. Madde ayırt ediciliği (MDISC) ($\alpha_1, \alpha_2, \alpha_3 \dots \alpha_k$) olmak üzere bir vektör ile temsil edilmektedir. Bir vektörün uzunluğu,

$$MDISC = \sqrt{\sum_{n=1}^k \alpha^2_{ik}} \quad (11)$$

ile ifade edilmektedir. Uzunluk, maddeye ait ortak ayırt edicilik olarak adlandırılmaktadır (Göçer Şahin, 2016). Uzunluk arttıkça maddeye ilişkin ayırt ediciliğin de arttığı söylenebilir. Yukarıdaki formülde ifade edilen α_{ik} , her bir boyuta ait ayırt edicilik değerlerini ifade etmektedir. Buradaki MDISC değeri de tekboyutlu MTK'deki ayırt edicilik gibi yorumlanabilir.

Bir vektörün uzunluğunun yanı sıra yönü ve başlangıç noktasına olan uzaklığının bilinmesinde yarar vardır. Bir vektörün yönü:

$$\alpha_i = \arccos\left(\frac{\alpha_{i1}}{MDISC}\right) \quad (12)$$

formülü ile ifade edilir. Formülde ifade edilen α_i madde vektörünün θ_1 eksenine yaptığı açıdır. Buna göre 45° lik açı, maddenin her iki yeteneği de iyi ölçtüğü anlamına gelir. Eğer açı 45° den büyükse, ikinci boyutun, birinci boyuttan daha iyi ölçüldüğü anlamına gelmektedir. Bununla birlikte 45° den küçük ise bu maddenin birincil olarak θ_1 yeteneğini ölçtüğü yani birinci boyutu ikinci boyuttan daha iyi ölçüldüğü anlamına gelmektedir (Göçer Şahin, 2016; Sünbül, 2011).

Tek boyutlu MTK'de b parametresinin ÇBMTK'deki karşılığı, madde vektörünün başlangıç noktasına olan uzaklığını ifade eden D parametresi madde güçlüğü hakkında bilgi verir (Reckase, 2009). Bu parametre,

$$D = \frac{-d_i}{MDISC} \quad (13)$$

formülü ile hesaplanmaktadır. Formülde yer alan d_i , ara kesit terimi olarak nitelendirilir. D parametresi, tekboyutlu madde tepki kuramındaki b parametresine karşılık gelmektedir. Negatif işaretli olması maddenin kolay; pozitif olması ise maddenin zor olduğu şeklinde yorumlanmaktadır.

Bu araştırmada, çokboyutlu maddeler için MDISC parametresinin ranjı 0.8 ve 1.8 olarak girilmiştir. Bu ranjin belirlenmesinde Ackerman'ın yaptığı çalışma göz önünde bulundurulmuştur. Basit madde sayısının 10 ve yapının 2 boyutlu olduğu koşulda, maddenin yapısı, parametreler ve maddelerin eksenlerle yaptığı açılar örnek teşkil etmesi amacıyla Çizelge 6'da sunulmuştur:

Çizelge 6.

Madde parametreleri

<i>Boyutlar</i>	<i>Maddeler</i>	a_{j1}	a_{j2}	b	<i>MDISC</i>	D	<i>Açı</i>
	1	1.265	.111	-.579	1.27	.46	5.00
	2	1.074	.126	.422	1.08	-.39	6.67
	3	1.671	.245	-.109	1.69	.06	8.33
	4	1.312	.231	-.533	1.33	.40	10.00
	5	.980	.202	-.233	1.00	.23	11.67
	6	.937	.222	-.123	.96	.13	13.33
	7	.903	.242	-.726	.93	.78	15.00
	8	1.164	.349	.415	1.22	-.34	16.67
	9	1.076	.356	.074	1.13	-.07	18.33
	10	.765	.278	-.147	.81	.18	20.00
	11	.434	1.194	-.579	1.27	.46	70.00
	12	.334	1.029	.422	1.08	-.39	72.00
	13	.465	1.623	-.109	1.69	.06	74.00
	14	.322	1.293	-.533	1.33	.40	76.00
	15	.208	.979	-.233	1.00	.23	78.00
	16	.167	.948	-.123	.96	.13	80.00
	17	.130	.926	-.726	.93	.78	82.00
	18	.127	1.209	.415	1.22	-.34	84.00
	19	.079	1.130	.074	1.13	-.07	86.00
	20	.028	.813	-.147	.81	.18	88.00

2.3.Verilerin Analizi

Verilerin analizi aşağıdaki adımlarla gerçekleştirilmiştir:

1. Tekboyutluluğun değerlendirilmesinde farklı yöntemlerin performanslarını karşılaştırmak amacıyla hem parametrik hem de parametrik olmayan yöntemlerden yararlanılmıştır. Bu araştırma kapsamında parametrik olmayan yöntemlerden DIMTEST T istatistiği, Boyutluluk DETECT IDN indeksi kullanılmıştır. Parametrik yöntemlerden ise Doğrusal Olmayan Faktör Analizi (NOHARM) yönteminden yararlanılmıştır.
2. 1. tip hataları ve güç oranlarının değerlendirilmesinde sırasıyla tekboyutlu ve çok boyutlu veriler oluşturulmuştur.
3. Stout ve diğ.(1996), Forelich (2006) çalışmalarının yanı sıra Forelich ve Habing (2008) DIMTEST T istatistiği için AT ve PT bölümlenmesi yaparken, (a) AT maddelerinin boyutluluk açısından homojen olması gerektiği bir başka ifadeyle geometrik temsilde bunun anlamı, AT maddelerinin bulunduğu açının görece dar olması gerektiği ifade edilmiştir. (b) Θ_{AT} ile Θ_{PT} olabildiğince farklı olmalıdır, bir başka ifadeyle geometrik temsilde bunun anlamı, Θ_{AT} ile Θ_{PT} arasındaki açılar mümkün olduğunca geniş olmalıdır. (c) AT'nin içinde en az dört madde, PT'de ise testte yer alan maddelerin en az yarısı bulunmalıdır. Bu araştırmada, DIMTEST T istatistiği için AT ve PT maddeleri madde sayısının yarısı AT alt testinde diğer yarısı da PT alt testinde olacak şekilde tüm koşullar için sabitlenmiştir.
4. DIMTEST T istatistiği kritik 1,96 değerinden büyük çıktığı durumlarda H_0 hipotezi reddedilmiştir.
5. Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi yöntemleri varsayılan seçenekleriyle kullanılmıştır. Boyutluluk DETECT IDN indeksi değerinin 1 ya da daha büyük olması çok boyutluluğun yüksek olduğunun, 0.4 ile 1 arası olması çok boyutluluğun orta derecede olduğunun ve 0.2 ile 0.4 arasında olması ise tekboyutluluğun göstergesidir. Kim (1994) yaptığı simülasyon çalışmasında Boyutluluk DETECT IDN indeksi 0.10'dan küçük

olduğu durumda verilerin tekboyutlu olduğu kabul edilebileceğini, 0.10 ile 0.50 arasında olması çok boyutluluğun zayıf bir ihtimal olabileceği, 0.51 ile 1 arasında olması çok boyutluluğun orta düzeyde olduğunu ve 1'in üzerinde çıkması güçlü çok boyutluluğa işaret ettiğini ifade etmiştir (Ackerman ve Walker, 2003).

6. Tüm analizler için 100 kere replikasyon gerçekleştirilmiştir. DIMTEST T istatistiği ve Boyutluluk DETECT IDN indeksinin her bir koşulu için güvenilirlik katsayıları, üretilen veriler, theta değerleri, DIMTEST T istatistiği ve Boyutluluk DETECT IDN indeksi olmak üzere 5 farklı sonuç tablosu elde edilmiştir. DIMTEST T istatistiği için T istatistiği ve p anlamlılık değerleri rapor edilmiştir. DETECT istatistiği için ise DETECT değeri ve IDN indeksi rapor edilmiştir.
7. Parametrik yöntemlerden doğrusal olmayan faktör analizinde de (NOHARM) güvenilirlik katsayıları, üretilen veriler, theta değerleri ve NOHARM sonuç tabloları elde edilmiştir. NOHARM programının çıktılarının yorumlanmasında Tanaka Uyum İyiliği İndeksi (TIGF) ve RMSR olmak üzere iki indeksten yararlanılmıştır. TIGF değerinin ≥ 0.95 çıkması ve RMSR değerinin ≤ 0.05 çıkması modelin iyi uyum gösterdiğinin kanıtıdır (Hooper ve ark., 2008, Hu ve Bentler, 1999).
8. Tüm sonuçlar için tekboyutluluk reddetme oranları her bir koşul için rapor edilmiştir.

BÖLÜM III

3. BULGULAR ve YORUMLAR

Bu bölümde verilerin analiz edilmesi ile elde edilen bulgulara ve bu bulguların alan yazın ile ilişkisine dair yorumlara araştırma sorularının sırası dikkate alınarak yer verilmiştir.

3.1. "1-0 Puanlanan Testlerde Çeşitli Boyutluluk Belirleme Yöntemlerine Göre 1. Tip Hata Oranı Nasıl Değişmektedir?" Bir Başka İfadeyle 1-0 Puanlanan Testlerde Çeşitli Boyutluluk Belirleme Yöntemlerine Göre Testin Uzunluğunun, Dağılımın Özelliğinin ve Örneklem Büyüklüğünün Manipule Edildiği Tekboyutlu Verilerden Elde Edilen 1. Tip Hata Oranları Nasıl Değişmektedir?

Standart normal dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu ve testin uzunluğunun 10, 20 ve 30 madde olduğu koşullarda DIMTEST T istatistiği sonuçları Çizelge 7’de sunulmuştur:

Çizelge 7.

Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T İstatistiği için 1. tip hata oranları

Örneklem Büyüklüğü	Madde Sayısı	T Ortalama Değeri	Reddetme oranı
200	10	0.23	0.04
	20	-0.23	0.00
	30	0.22	0.03
300	10	0.45	0.06
	20	-0.09	0.00
	30	0.01	0.00
500	10	0.37	0.02
	20	0.12	0.00
	30	-0.03	0.00
1000	10	0.60	0.05
	20	0.05	0.00
	30	0.22	0.00
2000	10	1.04	0.18
	20	-0.09	0.00
	30	0.08	0.01
3000	10	0.84	0.10
	20	-0.30	0.00
	30	0.29	0.01

N(0,1): Standart Normal Dağılım, replikasyon sayısı:100, DIMTEST Boyutluluk T İstatistiği için kullanılan program: DIMTEST.

Çizelge 7’de yer alan DIMTEST T istatistiği sonuçlarına göre, örneklem büyüklüğü kaç olursa olsun tekboyutlu verilerde testin uzunluğu arttıkça tekboyutluluğu reddetme oranı genel olarak azalmaktadır. Bir başka ifadeyle, 1. tip hata oranı genel olarak azalmaktadır. Örneklem büyüklüğünün 300, 500 ve 1000 olduğu, testin uzunluğunun da 20 ve 30 madde olduğu durumda DIMTEST T istatistiği için 1. tip hata oranı 0.00 bulunmuştur. Bir başka ifadeyle bu koşullarda verilerin tekboyutlu olduğu sıfır hipotezi mükemmel bir şekilde kabul edilmiştir. Fakat bununla birlikte, testin uzunluğunun 10 madde olduğu durumda DIMTEST T istatistiği için 1. tip hata oranı 20 ve 30 madde olduğu duruma göre artmaktadır. Testin uzunluğunun 20 maddeden az olduğu durumlarda, DIMTEST T istatistiğinin standart normal dağılım gösteren verilerde tekboyutluluğu belirlemede hata ürettiği söylenebilir.

Seo ve Sünbül (2012) de yaptıkları çalışmada DIMTEST T istatistiğinin kısa testlerde tek boyutluluğu belirlerken kaydadeğer bir oranda 1. Tip hata oranı verdiğini ifade etmektedir. Bununla birlikte, DIMTEST T istatistiği tek başına kullanıldığında örneklem büyüklüğü, test uzunluğu ve AT ve PT olmak üzere iki alt test arasındaki madde sayıları oranına göre çeşitli ret oranları gösterdiği ifade edilmiştir. Yapılan çalışmada DIMTEST T istatistiğinin kısa testlerde örneklem büyüklüğünün yüksek olduğu durumlarda bile yüksek 1. tip hata oranları sergilediği ifade edilmiştir. Örneklem büyüklüğü arttıkça kısa testlerde 1.tip hata artışını önlemek için daha fazla PT alttest uzunluğu gerektiği ifade edilmiştir.

Örneklem büyüklüğü arttıkça DIMTEST T istatistiğinin –özellikle 20 ve 30 madde olduğu durumlarda- standart dağılım gösteren tekboyutlu veriler için doğru sonuçlar ürettiği görülmektedir. Özellikle 300 ve üzeri örneklem için daha doğru sonuçlar göstermiştir. Bu bulgunun Finch ve Habing (2007) ile Finch ve Monahan’ın (2008) çalışmalarını destekler nitelikte olduğu söylenebilir.

Socha ve DeMars (2013) ise, yaptıkları araştırmada testin uzunluğunun 20 madde ve örneklem büyüklüğünün 4000 olduğu koşul hariç yine testin uzunluğunun 20 madde fakat daha düşük örneklem büyüklüğüne sahip tüm koşullarda 1. tip hata oranının 0.05 oranının üstüne çıkmadığı ifade edilmiştir. Bu çalışmadaki bulgular Socha ve DeMars (2013)’ın bulgularıyla örtüşmekle birlikte testin uzunluğunun 20 madde olduğu koşulda örneklem büyüklüğünün yüksek olduğu durumlar için Socha ve DeMars (2013) ‘ın çalışması bu çalışma ve diğer çalışmalarla tutarsızlık göstermiştir.

Standart normal dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu ve testin uzunluğunun 10, 20 ve 30 madde olduğu koşullarda DETECT IDN indeksi için 1. Tip hata oranları Çizelge 8’de sunulmuştur:

Çizelge 8.

Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre Boyutluluk DETECT IDN indeksi için 1. tip hata oranları

Örneklem Büyüklüğü	Madde Sayısı	DETECT Değeri	IDN İndeksi	Reddetme oranı
200	10	0.03	0.52	0.71
	20	0.01	0.51	0.62
	30	0.00	0.50	0.53
300	10	0.05	0.51	0.64
	20	0.01	0.50	0.57
	30	0.01	0.50	0.38
500	10	0.05	0.51	0.59
	20	0.02	0.50	0.50
	30	0.02	0.50	0.38
1000	10	0.06	0.51	0.47
	20	0.01	0.51	0.43
	30	0.00	0.50	0.31
2000	10	0.09	0.52	0.39
	20	0.02	0.50	0.30
	30	0.01	0.51	0.20
3000	10	0.09	0.52	0.38
	20	0.02	0.51	0.28
	30	0.01	0.51	0.19

N(0,1): Standart Normal Dağılım, replikasyon sayısı:100, DETECT IDN indeksi için kullanılan program: DETECT.

Çizelge 8’de yer alan DETECT IDN indeksi sonuçlarına göre, standart normal dağılım gösteren verilerde, örneklem büyüklüğü arttıkça 1.tip hata oranı genel olarak azalmaktadır. Özellikle örneklem büyüklüğünün 200, 300 ve 500 olduğu durumda, 1. tip hata oranının yüksek olduğu görülmektedir. Bununla birlikte, testin uzunluğunun 10 madde olduğu durumlarda daha tutarsız sonuçlar verdiği görülmektedir. Roussos ve Özbek (2006) tarafından yapılan çalışmada, DETECT IDN indeksi sonuçlarının madde sayısının özellikle 10 ya da daha az olduğu ve veri setinin tekboyutlu olduğu durumlarda istatistiksel yanlışlık sergilediği belirtilmektedir. Buna göre araştırmacılar, 20 maddeden az test uzunluğu için DETECT IDN indeksini kullanmamayı önermişlerdir. Bu çalışma Roussos ve Özbek’in (2006) çalışmasıyla örtüşmekle birlikte

DETECT IDN indeksinin kullanılabilmesi için özellikle örneklem büyüklüğünün artırılması gerektiği göz önünde bulundurulması gereken önemli bir faktör olduğu bulgusuna ulaşılmıştır.

Standart normal dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu ve testin uzunluğunun 10, 20 ve 30 madde olduğu koşullarda doğrusal olmayan faktör analizi sonuçları için 1. Tip hata oranları Çizelge 9’da sunulmuştur:

Çizelge 9.

Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre Doğrusal Olmayan Faktör Analizi için 1. tip hata oranları

Örneklem Büyüküğü	Madde Sayısı	RMSR	TANAKA GFI	Reddetme oranı
200	10	0.00915	0.99221	0.00
	20	0.01082	0.98823	0.00
	30	0.01075	0.94888	0.00
300	10	0.00784	0.99451	0.00
	20	0.00904	0.99197	0.00
	30	0.00870	0.99024	0.00
500	10	0.00618	0.99651	0.00
	20	0.00678	0.99552	0.00
	30	0.00697	0.99627	0.00
1000	10	0.00433	0.99827	0.00
	20	0.00495	0.99756	0.00
	30	0.00482	0.99693	0.00
2000	10	0.0025	0.99946	0.00
	20	0.00334	0.99879	0.00
	30	0.00345	0.99836	0.00
3000	10	0.00257	0.99934	0.00
	20	0.00287	0.99919	0.00
	30	0.00288	0.99885	0.00

N(0,1): Standart Normal Dağılım, replikasyon sayısı:100, Doğrusal Olmayan Faktör Analizi için Kullanılan Program ve Elde Edilen İndeksler: NOHARM- RMSR ve TIGF

Çizelge 9’da yer alan standart normal dağılım gösteren tekboyutlu verilerde, boyutluluk belirleme parametrik yöntemlerinden doğrusal olmayan faktör analizi (NOHARM) sonuçları incelendiğinde, standart normal dağılım gösteren tekboyutlu veriler için, örneklem büyüklüğü ve testin uzunluğu kaç olursa olsun Tanaka Uyum İyiliği İndeksi (TIGF) değeri ≥ 0.95 çıkmıştır. Bununla beraber, tüm sonuçlarda RMSR değerinin ≤ 0.05 çıkması modelin iyi uyum gösterdiğinin kanıtı olmuştur. RMSR ve

Tanaka Uyum İyiliği İndeksi değerleri tekboyutluluk için 0.00 ret oranı sergilemiştir. Bir başka ifadeyle her koşulda verilerin tekboyutlu olduğu sıfır hipotezini doğru şekilde kabul etmiştir. Gessaroli ve De Champlain (1996) tarafından yapılan araştırmada, doğrusal olmayan faktör analizi (NOHARM)'nin iki parametrelili lojistik (2PL) modeller için 1. tip hata oranının için test uzunluğunun 15, 30 ve 45 olduğu koşullarda iyi performans gösterdiği ifade edilmiştir.

Standart normal dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu ve testin uzunluğunun 10, 20 ve 30 madde olduğu koşullarda DIMTEST T istatistiği, DETECT IDN indeksi ve doğrusal olmayan faktör analizi için 1. Tip hata oranları Çizelge 10'da sunulmuştur:

Çizelge 10.

Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T İstatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için 1. tip hata oranları

Örneklem Büyüküğü	Madde Sayısı	DIMTEST T İSTATİSTİĞİ	DETECT IDN İNDEKSİ	RMSR	TIGF
		Reddetme oranı	Reddetme oranı	Reddetme oranı	Reddetme oranı
200	10	0.04	0.71	0.00	0.00
	20	0.00	0.62	0.00	0.00
	30	0.03	0.53	0.00	0.00
300	10	0.06	0.64	0.00	0.00
	20	0.00	0.57	0.00	0.00
	30	0.00	0.38	0.00	0.00
500	10	0.02	0.59	0.00	0.00
	20	0.00	0.50	0.00	0.00
	30	0.00	0.38	0.00	0.00
1000	10	0.05	0.47	0.00	0.00
	20	0.00	0.43	0.00	0.00
	30	0.00	0.31	0.00	0.00
2000	10	0.18	0.39	0.00	0.00
	20	0.00	0.30	0.00	0.00
	30	0.01	0.20	0.00	0.00
3000	10	0.10	0.38	0.00	0.00
	20	0.00	0.28	0.00	0.00
	30	0.01	0.19	0.00	0.00

N(0,1): Standart Normal Dağılım, replikasyon sayısı:100, Boyutluluk T İstatistiği için kullanılan program: DIMTEST, DETECT IDN indeksi için kullanılan program: DETECT, Doğrusal Olmayan Faktör Analizi için Kullanılan Program ve Elde Edilen İndeksler: NOHARM- RMSR ve TIGF

Çizelge 10'da yer alan standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T istatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi sonuçları 1. tip hata oranları bakımından incelendiğinde, DIMTEST T istatistiği sonuçlarına göre, test uzunluğunun 10 madde olduğu durumda tekboyutluluk kararını vermede, test uzunluğunun 20 ve 30 madde olduğu durumlara göre daha tutarsız olduğu görülmektedir. Bunun yanı sıra test uzunluğunun 20 ve 30 olduğu durumda örneklem büyüklüğü kaç olursa olsun daha tutarlı sonuçlar verdiği görülmektedir. DIMTEST T istatistiği sonuçlarına göre, örneklem büyüklüğü kaç olursa olsun tekboyutlu verilerde testin uzunluğu arttıkça tekboyutluluğu reddetme oranının genel olarak azaldığı bir başka ifadeyle, 1. tip hata oranının azaldığı görülmektedir. Örneklem büyüklüğünün 300, 500 ve 1000 olduğu, testin uzunluğunun da 20 ve 30 madde olduğu durumda DIMTEST T istatistiği için 1. tip hata oranı 0.00 bulunmuştur. Diğer bir ifadeyle bu koşullarda verilerin tekboyutlu olduğu sıfır hipotezini mükemmel bir şekilde kabul etmiştir.

DIMTEST T istatistiği sonuçlarında bir diğer dikkat çeken nokta ise örneklem büyüklüğü arttıkça testin uzunluğunun özellikle 20 ve 30 madde olduğu durumlarda standart normal dağılım gösteren tekboyutlu veriler için doğru sonuçlar ürettiğidir. Özellikle 300 ve üzeri örneklem büyüklüğünde daha doğru sonuçlar göstermektedir. Bu bulgu Finch ve Habing (2007) ile Finch ve Monahan'ın (2008) çalışmalarını destekler nitelikte olduğu söylenebilir.

DETECT IDN indeksi sonuçları incelendiğinde, standart normal dağılım gösteren veriler için, örneklem büyüklüğü azaldıkça 1. tip hata oranı genel olarak artmaktadır. Özellikle örneklem büyüklüğünün 200, 300 ve 500 olduğu durumda, 1. tip hata oranının yüksek olduğu görülmektedir. Bununla birlikte, testin uzunluğunun 10 madde olduğu durumlarda daha tutarsız sonuçlar verdiği söylenebilir. Roussos ve Özbek (2006) tarafından yapılan çalışmada, DETECT IDN indeksinin özellikle 10 ya da daha az test uzunluğu olduğu ve verinin tekboyutlu olduğu durumlarda istatistiksel yanlışlık sergilediği ifade edilmiştir. Buna göre araştırmacılar, 20 maddeden az test uzunluğu için DETECT kullanmamayı önermişlerdir. Bu çalışma Roussos ve Özbek'in (2006) çalışmasıyla örtüşmekle birlikte DETECT yönteminin kullanılabilmesi için özellikle örneklem büyüklüğünün artırılması gerektiği göz önünde bulundurulması gereken önemli bir faktör olduğu bulgusuna ulaşılmıştır.

Parametrik boyutluluk belirleme yöntemlerinden doğrusal olmayan faktör analizi sonucu elde edilen RMSR ve Tanaka Uyum İyiliği İndeksleri incelendiğinde ise, standart normal dağılım gösteren tekboyutlu veriler için, örneklem büyüklüğü ve testin uzunluğu kaç olursa olsun Tanaka Uyum İyiliği İndeksi (TIGF) değeri ≥ 0.95 çıkmıştır. Bununla beraber, tüm sonuçlarda RMSR değerinin ≤ 0.05 çıkması modelin iyi uyum gösterdiğinin kanıtı olmuştur. RMSR ve Tanaka Uyum İyiliği İndeksi değerleri tekboyutluluk için 0.00 ret oranı sergilemiştir. Bir başka ifadeyle verilerin her koşulda tekboyutlu olduğu sıfır hipotezini doğru şekilde kabul etmiştir. Bu bulgunun, Twu (2005) ile Seo ve Sünbül'ün (2012) araştırma bulgularıyla örtüştüğü görülmektedir. Bununla birlikte Gessaroli ve De Champlain (1996) tarafından yapılan araştırmada, testin uzunluğunun 15, 30 ve 45 madde olduğu koşullarla da tutarlılık göstermiştir.

Çarpık dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu ve testin uzunluğunun 10, 20 ve 30 madde olduğu koşullarda DIMTEST T istatistiği için 1.tip hata oranları Çizelge 11'de sunulmuştur:

Çizelge 11.

Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T İstatistiği için 1. tip hata oranları

Örneklem Büyüklüğü	Madde Sayısı	T Ortalama Değeri	Reddetme oranı
200	10	0.32	0.04
	20	-0.21	0.00
	30	-0.15	0.00
300	10	0.65	0.05
	20	0.04	0.03
	30	-0.19	0.00
500	10	0.25	0.04
	20	-0.10	0.02
	30	-0.01	0.01
1000	10	0.73	0.09
	20	-0.18	0.00
	30	0.13	0.02
2000	10	0.78	0.16
	20	-0.02	0.01
	30	-0.11	0.00
3000	10	0.80	0.18
	20	-0.07	0.01
	30	0.02	0.01

(1.75, 3.75) Çarpık Dağılım, replikasyon sayısı:100, DIMTEST Boyutluluk T İstatistiği için kullanılan program: DIMTEST.

Çizelge 11’de yer alan DIMTEST T istatistiği sonuçlarına göre 1.tip hata oranları incelendiğinde, çarpık dağılım gösteren verilerde örneklem büyüklüğünün 200, 300, 500 ve 1000 olduğu durumda testin uzunluğu arttıkça tekboyutluluğu reddetme oranı azalmaktadır. Bir başka ifadeyle, 1. tip hata oranı azalmaktadır. Alexandra, Abswoude, Ark ve Sijitma (2004) tarafından, parametrik olmayan IRT modellerinde, DETECT, DIMTEST ve HCA/CCPROX yöntemlerinin karşılaştırıldığı çalışmada, DIMTEST ve DETECT’i temel alan istatistiklerin diğer istatistiklere göre tekboyutluluğu kabul ya da reddetmedeki doğruluğunun dikkat çekici olduğu ifade edilmiştir. Bu bulgunun Alexandra, Abswoude, Ark ve Sijitma’nın (2004) çalışmasıyla uyum içinde olduğu söylenebilir.

DIMTEST T istatistiği sonuçlarına göre, örneklem büyüklüğü arttıkça testin uzunluğunun 10 madde olduğu durumlar için çarpık dağılım gösteren verilerde 1. tip hata oranının genel olarak arttığı söylenebilir. Diğer bir ifadeyle, DIMTEST T istatistiğinin çarpık dağılım gösteren kısa test uzunluğuna sahip verilerde tekboyutluluğu belirlerken kaydadeğer bir oranda 1. tip hata oranı verdiği söylenebilir. DIMTEST T istatistiği sonuçlarına göre, testin uzunluğu arttıkça 1. tip hata oranı azalsa da standart normal dağılım gösteren verilere oranla çarpık dağılım gösteren verilerde azda olsa 1. tip hata oranlarının daha yüksek çıktığı görülmektedir.

Seraphine (2000), yetenek dağılımının normal olmadığı durumlarda DIMTEST’in performansını inceleyen araştırmasında, örtük özelliğin eğriliğinin 1. tip hata oranı ya da DIMTEST T istatistiğinin gücü üzerinde herhangi bir etkisinin olmadığını vurgulamıştır.

Çarpık dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu ve testin uzunluğunun 10, 20 ve 30 madde olduğu koşullarda DETECT IDN indeksi için 1. tip hata oranları sonuçları Çizelge 12’de sunulmuştur:

Çizelge 12.

Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre Boyutluluk DETECT IDN indeksi için 1. tip hata oranları

Örneklem Büyüklüğü	Madde Sayısı	DETECT Değeri	IDN indeksi	Reddetme oranı
200	10	0.48	-0.07	0.40
	20	0.48	-0.05	0.30
	30	0.48	-0.05	0.31
300	10	0.48	-0.09	0.39
	20	0.48	-0.07	0.21
	30	0.48	-0.05	0.29
500	10	0.47	-0.12	0.29
	20	0.46	-0.08	0.16
	30	0.42	-0.06	0.18
1000	10	0.47	-0.14	0.27
	20	0.46	-0.12	0.03
	30	0.47	-0.09	0.08
2000	10	0.45	-0.14	0.17
	20	0.44	-0.18	0.01
	30	0.46	-0.12	0.00
3000	10	0.42	-0.27	0.06
	20	0.43	-0.21	0.00
	30	0.45	-0.16	0.00

(1.75, 3.75) Çarpık Dağılım, replikasyon sayısı:100, DETECT IDN indeksi için kullanılan program: DETECT.

Çizelge 12’de yer alan çarpık dağılım gösteren veriler için DETECT IDN indeksi sonuçları incelendiğinde, test uzunluğunun 10 madde olduğu durumlarda 1. tip hata oranının her ne kadar örneklem büyüklüğü arttıkça azalsa da yine de yüksek olduğu görülmektedir. Özellikle örneklem büyüklüğünün 200, 300 ve 500 olduğu durumlarda testin uzunluğu kaç olursa olsun 1. tip hata oranının yüksek olduğu söylenebilir. Çarpık dağılım gösteren verilerde, tek boyutluluk belirlenirken DETECT IDN indeksi sonucunun en başarılı olduğu durumun örneklem büyüklüğünün 1000 ve üzeri, test uzunluğunun ise 20 ve 30 olduğu koşullarda olduğu görülmektedir. Bir başka ifadeyle örneklem büyüklüğü ve test uzunluğu arttıkça 1. tip hata oranının azaldığı söylenebilir.

DETECT IDN indeksinin kullanıldığı durumda, çarpık dağılım gösteren verilerde standart normal dağılım gösteren verilere göre sonuçların nispeten daha tutarlı olduğu ve 1. tip hata oranının daha düşük çıktığı görülmektedir.

Çarpık dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu ve testin uzunluğunun 10, 20 ve 30 madde olduğu koşullarda doğrusal olmayan faktör analizi için 1. Tip hata oranları Çizelge 13’de sunulmuştur:

Çizelge 13.

Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre Doğrusal Olmayan Faktör Analizi için 1. tip hata oranları

Örneklem Büyüklüğü	Madde Sayısı	RMSR	TANAKA GFI	Reddetme oranı
200	10	0.00902	0.99205	0.00
	20	0.01098	0.98618	0.00
	30	0.01081	0.98260	0.00
300	10	0.00742	0.99442	0.00
	20	0.00862	0.99180	0.00
	30	0.00873	0.98874	0.00
500	10	0.00665	0.99575	0.00
	20	0.00683	0.99487	0.00
	30	0.00689	0.99296	0.00
1000	10	0.00440	0.99818	0.00
	20	0.00496	0.99729	0.00
	30	0.00494	0.00637	0.00
2000	10	0.00280	0.00280	0.00
	20	0.02591	0.87180	0.00
	30	0.00357	0.99808	0.00
3000	10	0.00281	0.99850	0.00
	20	0.00293	0.99903	0.00
	30	0.00296	0.99868	0.00

(1.75, 3.75) Çarpık Dağılım, replikasyon sayısı:100, Doğrusal Olmayan Faktör Analizi için Kullanılan Program ve Elde Edilen İndeksler: NOHARM- RMSR ve TIGF

Çarpık dağılım gösteren verilerde, tek boyutluluğu belirlemek için kullanılan parametrik yöntemlerden doğrusal olmayan faktör analizi (NOHARM) sonuçlarına göre, çarpık dağılım gösteren tek boyutlu veriler için, örneklem büyüklüğü ve testin uzunluğu kaç olursa olsun Tanaka Uyum İyiliği İndeksi (TIGF) değeri ≥ 0.95 çıkmıştır. Bununla beraber, tüm sonuçlarda RMSR değerinin ≤ 0.05 çıkması modelin iyi uyum gösterdiğinin kanıtı olmuştur. RMSR ve Tanaka Uyum İndeksi değerleri tekboyutluluk için 0.00 ret oranı sergilemiştir. Bir başka ifadeyle her koşulda tekboyutlu olduğu sıfır hipotezi doğru olarak kabul edilmiştir. Doğrusal olmayan faktör analizi sonuçları, standart normal dağılım ve çarpık dağılım gösteren veriler için benzer sonuçlar ürettiği görülmüştür.

Çarpık dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu ve testin uzunluğunun 10, 20 ve 30 madde olduğu koşullarda DIMTEST T istatistiği, DETECT IDN indeksi ve doğrusal olmayan faktör analizi için 1. Tip hata oranları Çizelge 14’de sunulmuştur:

Çizelge 14.

Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T İstatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için 1. tip hata oranları

Örneklem Büyüklüğü	Madde Sayısı	DETECT IDN İNDEKSİ	DIMTEST T İSTATİSTİĞİ	NOHARM RMSR	NOHARM TIGF
		Reddetme oranı	Reddetme oranı	Reddetme oranı	Reddetme oranı
200	10	0.40	0.04	0.00	0.00
	20	0.30	0.00	0.00	0.00
	30	0.31	0.00	0.00	0.00
300	10	0.39	0.05	0.00	0.00
	20	0.21	0.03	0.00	0.00
	30	0.29	0.00	0.00	0.00
500	10	0.29	0.04	0.00	0.00
	20	0.16	0.02	0.00	0.00
	30	0.18	0.01	0.00	0.00
1000	10	0.27	0.09	0.00	0.00
	20	0.03	0.00	0.00	0.00
	30	0.08	0.02	0.00	0.00
2000	10	0.17	0.16	0.00	0.00
	20	0.01	0.01	0.00	0.00
	30	0.00	0.00	0.00	0.00
3000	10	0.06	0.18	0.00	0.00
	20	0.00	0.01	0.00	0.00
	30	0.00	0.01	0.00	0.00

(1.75, 3.75) Çarpık Dağılım, replikasyon sayısı:100, Boyutluluk T İstatistiği için kullanılan program:DIMTEST, DETECT IDN indeksi için kullanılan program: DETECT, Doğrusal Olmayan Faktör Analizi için Kullanılan Program ve Elde Edilen İndeksler: NOHARM- RMSR ve TIGF

Çizelge14’e göre, çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T istatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için 1. tip hata oranları incelendiğinde, DETECT IDN indeksi sonuçlarına göre, özellikle küçük örneklemelerde ve test uzunluğunun kısa olduğu durumlarda 1. tip hata oranının daha yüksek çıktığı görülmektedir. DIMTEST T istatistiğinin, DETECT IDN indeksine göre daha doğru

sonular verdiĐi grlmekle beraber testin uzunluĐunun 10 madde olduĐu kořulda DIMTEST T istatistiĐi sonularında, diĐer test uzunluklarına gre hata oranının daha yksek olduĐu bulunmuřtur. Bununla birlikte test uzunluĐunun 20 ve 30 madde olduĐu tm kořullarda DIMTEST T istatistiĐinin ok yksek oranda doĐru kararlar verdiĐi grlmektedir. DoĐrusal olmayan faktr analizi (NOHARM) sonuları incelendiĐinde ise, arpık daĐılım gsteren tekboyutlu veriler iin, rneklem byklĐ ve testin uzunluĐu ka olursa olsun 0.00 ret oranı sergilemiřtir. Bir bařka ifadeyle her kořulda verilerin tekboyutlu olduĐu sıfır hipotezini doĐru olarak kabul edilmiřtir.

3.2. “1-0 Puanlanan Testlerde eřitli Boyutluluk Belirleme Yntemlerine Gre Testin Gc Nasıl DeĐiřmektedir?” Bir Bařka İfadeyle 1-0 Puanlanan Testlerde eřitli Boyutluluk Belirleme Yntemlerine Gre Testin UzunluĐunun, Boyutlararası Korelasyon Derecesinin, DaĐılımların ve rneklem ByklĐnn Manipule EdildiĐi İki Boyutlu Veriler İin Testin Gc Nasıl DeĐiřmektedir?

Standart normal daĐılım gsteren verilerde, rneklem byklĐnn 200, 300, 500, 1000, 2000 ve 3000 olduĐu, testin uzunluĐunun 10, 20 ve 30 madde olduĐu ve boyutlararası korelasyonun derecesinin (0.25, 0.50, 0.75 ve 0.90) olduĐu kořullarda DIMTEST T istatistiĐi sonuları izelge 15’de sunulmuřtur:

Çizelge 15.

Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre DIMTEST T İstatistiği için güç oranları

N(0,1)		Boyutlar Arası Korelasyon							
		0.25		0.50		0.75		0.90	
Örneklem Büyüküğü	Madde Sayısı	T Ortalama Değeri	Reddetme oranı	T Ortalama Değeri	Reddetme oranı	T Ortalama Değeri	Reddetme oranı	T Ortalama Değeri	Reddetme oranı
200	10	2.50	0.91	1.96	0.64	1.30	0.37	0.70	0.14
	20	3.19	1.00	2.40	0.87	1.31	0.29	0.52	0.08
	30	3.59	1.00	2.86	0.95	1.73	0.52	0.91	0.13
300	10	3.39	0.99	2.62	0.95	1.79	0.55	2.11	0.26
	20	4.14	0.99	3.18	0.95	2.00	0.66	1.91	0.23
	30	4.45	1.00	3.55	1.00	2.33	0.86	1.25	0.27
500	10	4.17	1.00	3.20	0.97	1.98	0.69	1.37	0.30
	20	5.62	1.00	4.24	1.00	2.59	0.95	1.24	0.73
	30	6.03	1.00	4.60	1.00	2.96	0.98	1.58	0.47
1000	10	6.04	1.00	4.69	1.00	3.18	1.00	2.01	0.82
	20	7.78	1.00	5.74	1.00	3.37	1.00	1.59	0.56
	30	8.66	1.00	6.68	1.00	4.13	0.98	2.32	0.83
2000	10	8.98	1.00	6.80	1.00	4.49	1.00	2.90	0.96
	20	11.54	1.00	8.52	1.00	5.08	1.00	2.54	0.95
	30	12.42	1.00	9.72	1.00	5.94	1.00	3.09	0.98
3000	10	10.58	1.00	7.91	1.00	5.01	1.00	3.33	0.99
	20	14.01	1.00	10.40	1.00	6.06	1.00	3.06	0.98
	30	15.28	1.00	11.53	1.00	7.54	1.00	4.07	1.00

N(0,1): Standart Normal Dağılım, replikasyon sayısı:100 DIMTEST Boyutluluk T İstatistiği için kullanılan program: DIMTEST.

Çizelge 15'e göre DIMTEST T istatistiği sonuçları incelendiğinde, boyutlararası korelasyonun düşük (0.25), örneklem büyüklüğünün 200 ve testin uzunluğunun 10 madde olduğu koşul için iki boyutlu verilerde tek boyutluluğu reddetme oranı 0.91 iken örneklem büyüklüğünün 300 olduğu koşulda tek boyutluluğu reddetme oranı 0.99 olarak bulunmuştur. Boyutlararası korelasyonun 0.25 ve örneklem büyüklüğünün 500 ve üstü olduğu durumda testin uzunluğu kaç olursa olsun tüm koşullar için standart normal dağılım gösteren iki boyutlu verilerde tek boyutluluğu reddetme oranı 1.00 olarak bulunmuştur. Boyutlararası korelasyonun 0.50 olduğu, örneklem büyüklüğünün 500 ve testin uzunluğunun 10 olduğu koşul hariç tümünde tek boyutluluğu reddetme oranı 1.00 olarak bulunmuştur. Bunun yanı sıra örneklem büyüklüğünün 500'den az olduğu iki koşul için (200 ve 300) bir başka ifadeyle örneklem büyüklüğünün küçük olduğu durumda tek boyutluluğu reddetme oranı örneklem büyüklüğünün büyük olduğu durumlara göre daha düşük çıkmıştır. DIMTEST T istatistiğinin büyük örneklemli koşullarda, çokboyutluluk kararı için daha doğru sonuçlar verdiği söylenebilir. Bu bulgu Finch ve Habing (2007), Finch ve Monahan (2008) ve Özbek Baştuğ'un (2012) çalışmalarıyla tutarlılık göstermiştir. Örneklem büyüklüğünün özellikle 300'den daha az olduğu koşullar için DIMTEST T istatistiğinin yanılma oranı belirgin bir şekilde artmaktadır.

Boyutlararası korelasyonun 0.75 olduğu durumda, örneklem büyüklüğü arttıkça tek boyutluluğu reddetme oranı artmaktadır ancak örneklem büyüklüğünün 200 ve 300 olduğu durumda tekboyutluluğu reddetme oranı çok düşmektedir. Bir başka ifadeyle iki boyutlu veriyi büyük ölçüde tekboyutlu olarak kabul etmektedir. Boyutlararası korelasyonun 0.90 olduğu durumda ise, örneklem büyüklüğünün 200, 300 ve 500 olduğu koşul için iki boyutlu verilerde tekboyutluluk büyük oranda kabul edilirken örneklem büyüklüğü arttıkça tekboyutluluğu reddetme oranı belirgin bir şekilde artış göstermiştir.

Sonuç olarak DIMTEST T istatistiği sonuçlarına göre boyutlararası korelasyon değeri arttıkça iki boyutlu verilerde tekboyutluluğu reddetme oranı azalırken bir başka ifadeyle testin gücü azalırken bununla birlikte boyutlararası korelasyonun yüksek olduğu durumda örneklem büyüklüğü arttıkça tekboyutluluğu reddetme oranı artmaktadır. Boyutlararası korelasyonun düşük olduğu durumda ise örneklem büyüklüğü ve testin uzunluğu kaç olursa olsun tekboyutluluğu reddetme oranı DIMTEST T istatistiği sonuçları için 1.00'dir. Bir başka ifadeyle verinin iki boyutlu

olduğu kabul edilip, testin gücünün yüksek olduğu söylenebilir. DIMTEST T istatistiğinin çokboyutluluk kararı için, boyutlararası korelasyondan belirgin bir şekilde etkilendiği söylenebilir. Bununla birlikte örneklem büyüklüğü 3000 ve test uzunluğunun 30 olduğu durumda boyutlararası korelasyon değeri kaç olursa olsun 1.00 ret oranına ulaşmıştır. İlgili alanyazına bakıldığında, Zhang (2008) tarafından yapılan çalışmada boyutlar arası korelasyonun düşük olduğu koşul için, kısa testler uzun testlere göre daha iyi sonuçlar ürettiği ifade edilmiştir. Oysaki bu çalışmada boyutlararası korelasyonun çok düşük olduğu durumda test uzunluğu arttıkça testin gücü açısından DIMTEST T istatistiği sonuçları mükemmel performans göstermiştir. Zhang'ın (2008) çalışmasının aksine Alexandra, Abswoude, Ark ve Sijitma (2004), Seo ve Sünbül (2012) ve Özbek Baştuğ (2012) çalışmalarıyla ise bu araştırmadaki bulguların örtüştüğü görülmektedir.

Standart normal dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu, testin uzunluğunun 10, 20 ve 30 madde olduğu ve boyutlararası korelasyonun derecesinin (0.25, 0.50, 0.75 ve 0.90) olduğu koşullarda DETECT IDN İndeksi sonuçları Çizelge 16'da sunulmuştur:

Çizelge 16.

Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre Boyutluluk DETECT IDN İndeksi için güç oranları

		Boyutlar Arası Korelasyon											
Örneklem Büyüküğü	Madde Sayısı	0.25			0.50			0.75			0.90		
		DETECT	IDN	Reddetme oranı	DETECT	IDN	Reddetme oranı	DETECT	IDN	Reddetme oranı	DETECT	IDN	Reddetme oranı
200	10	-0.25	0.43	0.86	-0.18	0.44	0.92	-0.13	0.46	0.92	-0.11	0.46	0.96
	20	-0.06	0.48	1.00	-0.04	0.48	1.00	-0.02	0.49	1.00	-0.01	0.49	1.00
	30	-0.03	0.49	1.00	-0.02	0.49	1.00	-0.01	0.49	1.00	-0.01	0.49	1.00
300	10	-0.25	0.41	0.74	-0.18	0.43	0.85	-0.12	0.44	0.91	-0.04	0.46	0.98
	20	-0.05	0.47	1.00	-0.04	0.48	1.00	-0.03	0.48	1.00	-0.03	0.48	1.00
	30	-0.03	0.48	1.00	-0.02	0.49	1.00	-0.01	0.49	1.00	-0.01	0.49	1.00
500	10	1.14	0.89	1.00	0.78	0.79	1.00	0.37	0.66	0.99	0.18	0.57	1.00
	20	1.11	0.88	1.00	0.74	0.79	1.00	0.35	0.65	1.00	0.14	0.57	1.00
	30	0.96	0.85	1.00	0.63	0.77	1.00	0.31	0.65	1.00	0.12	0.56	1.00
1000	10	1.16	0.94	1.00	0.78	0.87	1.00	0.41	0.72	1.00	0.18	0.60	1.00
	20	1.11	0.94	1.00	0.71	0.86	1.00	0.36	0.72	1.00	0.14	0.59	1.00
	30	0.99	0.93	1.00	0.65	0.85	1.00	0.32	0.70	1.00	0.13	0.59	1.00
2000	10	1.15	0.98	1.00	0.78	0.92	1.00	0.40	0.78	1.00	0.17	0.63	1.00
	20	1.12	0.98	1.00	0.74	0.93	1.00	0.36	0.78	1.00	0.14	0.62	1.00
	30	0.99	0.97	1.00	0.65	0.91	1.00	0.32	0.76	1.00	0.13	0.61	1.00
3000	10	1.16	0.98	1.00	0.78	0.94	1.00	0.40	0.80	1.00	0.19	0.66	1.00
	20	1.12	0.99	1.00	0.74	0.96	1.00	0.36	0.82	1.00	0.15	0.65	1.00
	30	0.99	0.99	1.00	0.65	0.95	1.00	0.33	0.81	1.00	0.13	0.64	1.00

N(0,1): Standart Normal Dağılım, replikasyon sayısı:100

Çizelge 16'da yer alan DETECT IDN İndeksi sonuçlarına göre testin gücü incelendiğinde, standart normal dağılım gösteren iki boyutlu veriler için, boyutlararası korelasyonun 0.25 olduğu koşulda, örneklem büyüklüğünün 500 ve üstü olduğu durumda, testin uzunluğu kaç olursa olsun bir koşul hariç diğer tüm koşullarda tekboyutluluk için 1.00 ret oranı elde edilmiştir. Bir başka ifadeyle her koşulda testin tek boyutlu olduğu sıfır hipotezini doğru şekilde reddetmiştir. Örneklem büyüklüğünün 500, test uzunluğunun 10 madde, boyutlararası korelasyonun 0.75 olduğu durumda ise ret oranı 0.99 olarak elde edilmiştir. Bir diğer ifadeyle sıfır hipotezini 0.99 oranında doğru şekilde reddetmiştir. Bunun yanı sıra örneklem büyüklüğünün 200 ve 300 olduğu durumlarda testin uzunluğu azaldıkça tek boyutluluğu reddetme oranı düşmektedir. Özellikle testin uzunluğunun 10 madde olduğu durumda tek boyutluluğu reddetme oranlarında düşüş gözlenmektedir. Diğer bir ifadeyle yanlış karar verme oranının arttığı gözlenmiştir. Sonuç olarak iki boyutlu verilerde örneklem büyüklüğünün 500 ve üstü olduğu durumlarda, tekboyutluluğu reddedip iki boyutluluğu kabul etmede diğer bir deyişle testin gücünü belirlemede DETECT IDN İndeksi istatistiğinin iyi çalıştığı söylenebilir.

Svetina (2011)'nın DETECT ve NOHARM boyutluluk belirleme yöntemlerini kullanarak telafisel ve telafisel olmayan çok boyutlu madde tepki modellerindeki boyutluluğun değerlendirilmesinde karmaşık yapıların etkisini araştırdığı çalışmasında, DETECT'in büyük örneklemelerde, % 30 veya daha az karmaşık verilerde ve boyutlar arası korelasyonun 0.60 veya daha az olduğu durumlarda daha doğru oranlar sağladığını ifade edilmiştir. Bu araştırmadaki bulguların, Svetina'nın (2011) çalışmasındaki bulgular ile örtüştüğü söylenebilir.

Roussos ve Özbek (2006) tarafından yapılan ve çok boyutlu verilerde büyük örneklemli koşullarda DETECT indeksinin performansı incelenen araştırmada DETECT IDN indeksinin tüm simülasyon koşulları için oldukça küçük yanlışlık gösterdiğini belirtmişlerdir ve bununla birlikte DETECT' IDN indeksinin yüksek doğruluk oranı sağladığını ifade etmişlerdir. En düşük sınıflandırma oranının ise 20 maddelik yaklaşık basit yapılı iki boyutlu koşullar için 0.91 doğruluk oranı sağlandığı ifade edilmiştir. İlgili çalışmanın bu çalışmadaki bulgular ile örtüştüğü görülmektedir.

Standart normal dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu, testin uzunluğunun 10, 20 ve 30 madde olduğu ve boyutlararası korelasyonun derecesinin (0.25, 0.50, 0.75 ve 0.90) olduğu koşullarda doğrusal olmayan faktör analizi (NOHARM) sonuçları Çizelge 17'de sunulmuştur:

Çizelge 17.

Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre Doğrusal Olmayan Faktör Analizi için güç oranları

N(0, 1)		Boyutlar Arası Korelasyon											
		0.25			0.50			0.75			0.90		
Örneklem Büyüküğü	Madde Sayısı	RMSR	TANAKA GFI	Reddetme oranı	RMSR	TANAKA GFI	Reddetme oranı	RMSR	TANAKA GFI	Reddetme oranı	RMSR	TANAKA GFI	Reddetme oranı
200	10	0.0087	0.9931	1.00	0.0099	0.9880	1.00	0.0099	0.9882	1.00	0.0086	0.9920	1.00
	20	0.0120	0.9723	1.00	0.0106	0.9784	1.00	0.0114	0.9819	1.00	0.0123	0.9768	1.00
	30	0.0121	0.9550	1.00	0.0134	0.9522	1.00	0.0119	0.9722	1.00	0.0115	0.9765	1.00
300	10	0.0082	0.9913	1.00	0.0080	0.9917	1.00	0.0068	0.9947	1.00	0.0115	0.9846	1.00
	20	0.0113	0.9743	1.00	0.0111	0.9791	1.00	0.0105	0.9830	1.00	0.0101	0.9834	1.00
	30	0.0103	0.9709	1.00	0.0103	0.9780	1.00	0.0099	0.9769	1.00	0.0092	0.9826	1.00
500	10	0.0059	0.9963	1.00	0.0057	0.9967	1.00	0.0052	0.9974	1.00	0.0053	0.9976	1.00
	20	0.0066	0.9920	1.00	0.0075	0.9887	1.00	0.006	0.9941	1.00	0.0078	0.9934	1.00
	30	0.0070	0.9980	1.00	0.0072	0.9891	1.00	0.0069	0.9904	1.00	0.0066	0.9923	1.00
1000	10	0.0041	0.9982	1.00	0.0042	0.9982	1.00	0.0030	0.9992	1.00	0.0040	0.9986	1.00
	20	0.0049	0.9956	1.00	0.0043	0.9960	1.00	0.0051	0.9962	1.00	0.0051	0.9961	1.00
	30	0.0052	0.9930	1.00	0.0051	0.9954	1.00	0.0050	0.9954	1.00	0.0046	0.9962	1.00
2000	10	0.0029	0.9991	1.00	0.0026	0.9993	1.00	0.0026	0.9994	1.00	0.0026	0.9994	1.00
	20	0.0032	0.9986	1.00	0.0033	0.9978	1.00	0.0033	0.9984	1.00	0.0033	0.9984	1.00
	30	0.0035	0.9970	1.00	0.0035	0.9974	1.00	0.0035	0.9976	1.00	0.0035	0.9979	1.00
3000	10	0.0028	0.9999	1.00	0.0023	0.9994	1.00	0.0021	0.9995	1.00	0.0028	0.9993	1.00
	20	0.0027	0.9985	1.00	0.0027	0.9988	1.00	0.0026	0.9989	1.00	0.0026	0.9990	1.00
	30	0.0031	0.9975	1.00	0.0029	0.9981	1.00	0.0028	0.9985	1.00	0.0029	0.9980	1.00

N(0,1): Standart Normal Dağılım, replikasyon sayısı:100

Standart normal dağılım gösteren verilerde, parametrik yöntemlerden doğrusal olmayan faktör analizi (NOHARM) sonuçlarına göre testin gücü incelendiğinde, standart normal dağılım gösteren iki boyutlu veriler için, boyutlararası korelasyon, örneklem büyüklüğü ve testin uzunluğu kaç olursa olsun Tanaka Uyum İyiliği İndeksi (TIGF) değeri ≥ 0.95 çıkmıştır. Bununla beraber, tüm sonuçlarda RMSR değerinin ≤ 0.05 çıkması modelin iyi uyum gösterdiğinin kanıtı olmuştur. RMSR ve Tanaka Uyum İndeksi değerleri tekboyutluluk için 1.00 ret oranı sergilemiştir. Bir başka ifadeyle her koşulda testin tekboyutlu olduğu sıfır hipotezi doğru şekilde reddedilmiştir. İlgili alan yazın incelendiğinde Kaya ve Kelecioğlu (2016) tarafından yapılan çalışmada Doğrusal olmayan faktör analizinin (NOHARM) çokboyutluluğu belirlemede 50 ve üzeri örneklemelerde daha tutarlı olduğu ifade edilmiştir. Bu bulgu, Kaya ve Kelecioğlu (2016)'nın çalışmasıyla örtüşmekle birlikte, bu çalışmanın aksine Özbek Baştuğ'un (2012) ve Seo ve Sünbül'ün (2012) yaptığı çalışmalarıyla ise tutarsızlık göstermiştir. Bu çalışmalarda, doğrusal olmayan faktör analizinin (NOHARM) çok boyutluluk belirlemede güçlü bir istatistik yöntem olmadığı bulgularına ulaşılmıştır. Bunda ilgili çalışmalardaki karmaşık test yapısının etkili olduğu düşünülmektedir. Svetina (2011) tarafından yapılan çalışmada ise telafisel olmayan çok boyutlu MTK modellerine uygun verilerde, boyutluluğu belirlemede NOHARM'ı temel alan istatistiklerin DETECT'e göre daha güçlü performans gösterdiğini ifade etmiştir. Svetina (2011)'nin bu bulgusu, bu araştırmayı destekler nitelikte olduğu söylenebilir.

Standart normal dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu, testin uzunluğunun 10, 20 ve 30 madde olduğu ve boyutlararası korelasyonun derecesinin (0.25, 0.50, 0.75 ve 0.90) olduğu koşullarda DIMTEST T istatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi (NOHARM) sonuçlarına için testin gücü oranları Çizelge 18'de sunulmuştur:

Çizelge 18.

Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre DIMTEST T İstatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için güç oranları

N~(0,1)		Boyutlar Arası Korelasyon															
		0.25				0.50				0.75				0.90			
		Reddetme Oranı				Reddetme Oranı				Reddetme Oranı				Reddetme Oranı			
Örneklem Büyükliği	Madde Sayısı	DIMTEST T İSTATİSTİĞİ	Boyutluluk DETECT IDN İNDEKSİ	RMSR	TIGF	DIMTEST T İSTATİSTİĞİ	Boyutluluk DETECT IDN İNDEKSİ	RMSR	TIGF	DIMTEST T İSTATİSTİĞİ	Boyutluluk DETECT IDN İNDEKSİ	RMSR	TIGF	DIMTEST T İSTATİSTİĞİ	Boyutluluk DETECT IDN İNDEKSİ	RMSR	TIGF
200	10	0.91	0.86	1.00	1.00	0.64	0.92	1.00	1.00	0.37	0.92	1.00	1.00	0.14	0.96	1.00	1.00
	20	1.00	1.00	1.00	1.00	0.87	1.00	1.00	1.00	0.29	1.00	1.00	1.00	0.08	1.00	1.00	1.00
	30	1.00	1.00	1.00	1.00	0.95	1.00	1.00	1.00	0.52	1.00	1.00	1.00	0.13	1.00	1.00	1.00
300	10	0.99	0.74	1.00	1.00	0.95	0.85	1.00	1.00	0.55	0.91	1.00	1.00	0.26	0.97	1.00	1.00
	20	0.99	1.00	1.00	1.00	0.95	1.00	1.00	1.00	0.66	1.00	1.00	1.00	0.23	1.00	1.00	1.00
	30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.86	1.00	1.00	1.00	0.27	1.00	1.00	1.00
500	10	1.00	1.00	1.00	1.00	0.97	1.00	1.00	1.00	0.69	0.99	1.00	1.00	0.30	1.00	1.00	1.00
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.95	1.00	1.00	1.00	0.73	1.00	1.00	1.00
	30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00	0.47	1.00	1.00	1.00
1.000	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.82	1.00	1.00	1.00
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.56	1.00	1.00	1.00
	30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00	0.83	1.00	1.00	1.00
2000	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.96	1.00	1.00	1.00
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.95	1.00	1.00	1.00
	30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00
3000	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00
	30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

N(0,1): Standart Normal Dağılım, replikasyon sayısı:100, Boyutluluk T İstatistiği için kullanılan program: DIMTEST, DETECT IDN indeksi için kullanılan program: DETECT, Doğrusal Olmayan Faktör Analizi için Kullanılan Program ve Elde Edilen İndeksler: NOHARM- RMSR ve TIGF

Çizelge 18’de yer alan sonuçlara göre, standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre DIMTEST T İstatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için güç oranları incelendiğinde, DIMTEST T istatistiği sonuçlarına göre, boyutlararası korelasyonun 0.25 ve örneklem büyüklüğünün 500 ve üstü olduğu durumda testin uzunluğu kaç olursa olsun tüm koşullar için standart normal dağılım gösteren iki boyutlu verilerde tekboyutluluğu reddetme oranı 1.00 olarak bulunmuştur. Boyutlararası korelasyonun 0.50 olduğu, örneklem büyüklüğünün 500 ve testin uzunluğunun 10 madde olduğu koşul hariç tüm koşullarda tekboyutluluğu reddetme oranı 1.00 olarak bulunmuştur. Bunun yanı sıra örneklem büyüklüğünün 500’den az olduğu iki koşul için (200 ve 300) tekboyutluluğu reddetme oranı örneklem büyüklüğünün daha büyük olduğu durumlara göre daha düşük çıkmıştır.

Sonuç olarak DIMTEST T istatistiğinin büyük örneklemli koşullarda daha doğru sonuçlar verdiği söylenebilir. Bu bulgu Finch ve Habing (2007), Finch ve Monahan (2008) ve Özbek Baştuğ’un (2012) çalışmalarıyla tutarlılık göstermiştir. Özellikle örneklem büyüklüğünün 300’den daha az olduğu durumlar için DIMTEST T istatistiğinin belirgin bir şekilde yanılma oranı artmaktadır. DIMTEST T istatistiği sonuçlarına göre, boyutlar arası korelasyon değeri arttıkça iki boyutlu verilerde tekboyutluluğu reddetme oranı azalırken bir başka ifadeyle testin gücü azalırken bununla birlikte boyutlar arası korelasyonun yüksek olduğu durumda örneklem büyüklüğü arttıkça tekboyutluluğu reddetme oranı artmaktadır. Boyutlar arası korelasyonun düşük olduğu durumda ise örneklem büyüklüğü ve testin uzunluğu kaç olursa olsun tekboyutluluğu reddetme oranı DIMTEST T istatistiği için 1.00’dir. Bir başka ifadeyle verinin iki boyutlu olduğu kabul edilip, testin gücünün yüksek olduğu görülmektedir. DIMTEST T istatistiğinin çokboyutluluk kararı için, boyutlar arası korelasyondan belirgin bir şekilde etkilendiği görülmektedir. Bununla birlikte örneklem büyüklüğü 3000 ve test uzunluğunun 30 olduğu durumda boyutlar arası korelasyon değerine bakılmaksızın tekboyutluluk için 1.00 ret oranına ulaşmıştır. Diğer bir ifadeyle çokboyutluluk için mükemmel bir şekilde doğru karar verilmiştir. Zhang (2008) tarafından yapılan çalışmada boyutlar arası korelasyonun düşük olduğu koşulda, kısa testler uzun testlere göre daha iyi sonuçlar ürettiği ifade edilmiştir. Oysaki bu çalışmada boyutlar arası korelasyonun çok düşük olduğu durumda test uzunluğu arttıkça testin gücü açısından DIMTEST T istatistiği sonuçları mükemmel performans göstermiştir.

Araştırmanın bu sonucu Zhang'ın (2008) çalışmasıyla tutarsızlık gösterse de Alexandra, Abswoude, Ark ve Sijitma (2004), Seo ve Sünbül (2012) ve Özbek Baştuğ'un (2012) çalışmalarıyla örtüştüğü görülmektedir.

Testin gücü için Boyutluluk DETECT IDN indeksi sonuçları incelendiğinde, standart normal dağılım gösteren iki boyutlu veriler için, örneklem büyüklüğü 500 ve üstü olduğu durumda, boyutlar arası korelasyon değeri ve testin uzunluğu kaç olursa olsun, bir koşul hariç diğer tüm koşullar için 1.00 ret oranı sergilemiştir. Bir başka ifadeyle her koşulda tek boyutlu olduğu sıfır hipotezini doğru şekilde reddetmiştir. İki boyutlu verilerde örneklem büyüklüğünün 500 ve üstü olduğu durumlarda, tek boyutluluğu reddedip iki boyutluluğu kabul etmede Boyutluluk DETECT IDN istatistiğinin iyi çalıştığı söylenebilir. Bu bulgu Svetina'nın (2011) araştırmasındaki bulgularla ve Roussos ve Özbek'in (2006) çalışmalarıyla örtüştüğü görülmektedir.

Standart normal dağılım gösteren verilerde testin gücü için parametrik yöntemlerden doğrusal olmayan faktör analizi (NOHARM) sonucu elde edilen RMSR ve Tanaka Uyum İyiliği İndeksi değerleri incelendiğinde, standart normal dağılım gösteren iki boyutlu veriler için, boyutlar arası korelasyon, örneklem büyüklüğü ve testin uzunluğu kaç olursa olsun tek boyutluluk için 1.00 ret oranı sergilemiştir. Bir başka ifadeyle her koşulda testin tekboyutlu olduğu sıfır hipotezi doğru şekilde reddedilmiştir. İlgili alan yazın incelendiğinde Kaya ve Kelecioğlu (2016) tarafından yapılan çalışmada doğrusal olmayan faktör analizi sonuçlarının çokboyutluluğu belirlemede 50 ve üzeri örneklemelerde daha tutarlı olduğu ifade edilmiştir. Bu çalışmanın aksine Özbek Baştuğ'un (2012) ve Seo ve Sünbül'ün (2012) yaptığı çalışmada ise, doğrusal olmayan faktör analizinin (NOHARM) çokboyutluluk belirlemede güçlü bir istatistik yöntem olmadığı bulgularına ulaşılmıştır ancak bununla birlikte Svetina (2011) ise telafisel olmayan çokboyutlu MTK modellerine uygun verilerde, boyutluluğu belirlemede doğrusal olmayan faktör analizi (NOHARM) sonuçlarını temel alan istatistiklerin Boyutluluk DETECT IDN indeksine göre daha güçlü performans gösterdiğini ifade etmiştir.

Sonuç olarak standart normal dağılım gösteren verilerde, Boyutluluk DETECT IDN indeksi ve doğrusal olmayan faktör analizi (NOHARM) sonuçlarının tüm koşullarda DIMTEST T istatistiğine göre daha doğru kararlar verdiği görülmektedir. DIMTEST T istatistiğinin ise boyutlar arası korelasyonun düşük olduğu koşullarda ve örneklem büyüklüğünün yüksek olduğu durumlarda daha doğru kararlar verdiği

söylenbilir. Ayrıca boyutlar arası korelasyonun yüksek olduğu durumlarda DIMTEST T istatistiğinin 2000 ve üzeri örneklemelerde daha iyi çalıştığı ifade edilebilir.

Çarpık dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu, testin uzunluğunun 10, 20 ve 30 madde olduğu ve boyutlararası korelasyonun derecesinin (0.25, 0.50, 0.75 ve 0.90) olduğu koşullarda iki boyutlu veriler için DIMTEST T istatistiği sonuçları Çizelge 19’da sunulmuştur:

Çizelge 19.

Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre DIMTEST T İstatistiği için güç oranları

(1.75, 3.75)		Boyutlar Arası Korelasyon									
		0.25		0.50		0.75		0.90			
200	10	2.38	0.84	1.97	0.69	1.15	0.24	0.85	0.14		
	20	2.85	0.96	2.35	0.89	1.48	0.36	0.67	0.08		
	30	2.70	0.90	2.38	0.82	1.77	0.61	0.81	0.13		
300	10	3.05	0.98	2.67	0.91	1.93	0.61	1.21	0.26		
	20	3.91	1.00	3.22	1.00	3.21	1.00	1.05	0.23		
	30	4.06	1.00	3.56	1.00	2.49	0.90	1.26	0.27		
500	10	3.80	1.00	2.92	0.96	2.10	0.77	1.22	0.73		
	20	5.19	1.00	4.15	1.00	2.73	0.92	1.34	0.68		
	30	5.40	1.00	4.65	0.99	3.04	0.96	1.51	0.40		
1000	10	5.80	1.00	4.85	1.00	3.09	0.97	2.03	0.68		
	20	7.25	1.00	5.92	1.00	3.67	1.00	1.79	0.56		
	30	8.11	1.00	6.83	1.00	4.39	0.98	2.38	0.82		
2000	10	8.35	1.00	6.80	1.00	4.51	1.00	2.92	0.99		
	20	10.74	1.00	8.62	1.00	5.56	1.00	2.84	0.98		
	30	11.75	1.00	9.60	1.00	6.15	1.00	3.13	0.97		
3000	10	10.09	1.00	8.23	1.00	5.30	1.00	3.23	0.99		
	20	13.12	1.00	10.61	1.00	6.77	1.00	3.28	0.98		
	30	14.58	1.00	12.00	1.00	7.64	1.00	3.91	0.99		

(1.75, 3.75) Çarpık Dağılım, replikasyon sayısı:100

Çizelge19’da yer alan çarpık dağılım gösteren iki boyutlu veriler için DIMTEST T istatistiği sonuçları incelendiğinde, boyutlar arası korelasyonun düşük (0.25), örneklem büyüklüğünün 500 ve üstü olduğu durumda test uzunluğu kaç olursa olsun tekboyutluluğu reddetme oranı 1.00 olarak bulunmuştur. Bu sonuç standart normal dağılım gösteren verilerde DIMTEST T istatistiği sonuçları için de benzerdir. Ancak

örneklem büyüklüğünün 200 ve 300 olduğu durumda testin uzunluğu azaldıkça iki boyutlu verilerde tekboyutluluğu reddetme oranında düşüş gözlenmektedir. Diğer bir ifadeyle II.tip hata oranı artmaktadır. Bunun yanı sıra örneklem büyüklüğünün 200 ve 300 olduğu koşulda, çarpık dağılım gösteren verilerde, standart normal dağılım gösteren verilere göre yanlış karar verme oranı daha yüksektir.

Örneklem büyüklüğünün 200 ve 300 olduğu koşullarda, boyutlararası korelasyon değeri arttıkça iki boyutlu verilerde tekboyutluluk büyük oranda kabul edilmektedir. Boyutlararası korelasyonun 0.50 olduğu durumda, örneklem büyüklüğünün 500 ve testin uzunluğunun 10 ve 30 madde olduğu koşullar hariç tüm koşullarda tekboyutluluğu reddetme oranı 1.00 olarak bulunmuştur.

Boyutlararası korelasyonun 0.75 olduğu koşulda, örneklem büyüklüğü arttıkça tek boyutluluğu reddetme oranı artmaktadır. Bu durum standart normal dağılım gösteren verilerde benzer olmakla birlikte standart normal dağılım gösteren verilerde örneklem büyüklüğünün 1000 ve testin uzunluğunun 10 madde olduğu koşulda tek boyutluluğu reddetme oranı 1.00 bulunurken çarpık dağılım gösteren veriler için aynı koşullarında tek boyutluluğu reddetme oranı 0.97 olarak bulunmuştur. Boyutlararası korelasyonun 0.90 olduğu durumda ise, örneklem büyüklüğünün 500 olduğu koşul için testin uzunluğu arttıkça tekboyutluluğu reddetme oranı da azalmaktadır. Bununla birlikte aynı koşullarda standart normal dağılım gösteren verilere göre çokboyutluluğu reddetme oranının daha yüksek olduğu görülmektedir. Boyutlararası korelasyonun 0.90 olduğu durumda, örneklem büyüklüğü arttıkça tekboyutluluğu reddetme oranı belirgin bir şekilde artış göstermiştir.

Sonuç olarak boyutlararası korelasyon değeri arttıkça tekboyutluluğu reddetme oranı azalırken bununla birlikte boyutlararası korelasyonun yüksek olduğu koşulda örneklem büyüklüğü arttıkça tekboyutluluğu reddetme oranı artmaktadır. Boyutlararası korelasyonun düşük olduğu durumda ise örneklem büyüklüğü ve testin uzunluğu kaç olursa olsun tekboyutluluğu reddetme oranı DIMTEST T istatistiği sonuçları için 1.00'dir. Bir başka ifadeyle tüm replikasyonlarda tekboyutluluk reddedilmiştir. Çarpık dağılımlar için, DIMTEST T istatistiğinin çokboyutluluk kararı için, boyutlararası korelasyondan belirgin bir şekilde etkilendiği söylenebilir.

Çarpık dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu, testin uzunluğunun 10, 20 ve 30 madde olduğu ve boyutlararası korelasyonun derecesinin (0.25, 0.50, 0.75 ve 0.90) olduğu koşullarda Boyutluluk DETECT IDN istatistiği sonuçları Çizelge 20'de sunulmuştur:

Çizelge 20.

Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre Boyutluluk DETECT IDN İndeksi için güç oranları

		Boyutlar Arası Korelasyon											
		0.25			0.50			0.75			0.90		
Örneklem Büyüküğü	Madde Sayısı	DETECT	IDN	Reddetme oranı	DETECT	IDN	Reddetme oranı	DETECT	IDN	Reddetme oranı	DETECT	IDN	Reddetme oranı
200	10	-0.30	0.42	0.76	-0.26	0.42	0.79	-0.23	0.43	0.86	-0.21	0.43	0.85
	20	-0.11	0.46	1.00	-0.10	0.46	1.00	-0.10	0.46	0.99	-0.09	0.46	0.99
	30	-0.06	0.48	1.00	-0.05	0.48	1.00	-0.07	0.47	1.00	-0.05	0.48	1.00
300	10	-0.32	0.40	0.67	-0.27	0.41	0.71	-0.24	0.40	0.67	-0.23	0.41	0.77
	20	-0.11	0.46	1.00	-0.11	0.46	1.00	-0.10	0.46	1.00	-0.09	0.46	0.99
	30	-0.06	0.47	1.00	-0.06	0.47	1.00	-0.05	0.47	1.00	-0.05	0.47	1.00
500	10	1.19	0.89	1.00	0.84	0.81	1.00	0.47	0.70	1.00	0.24	0.60	0.99
	20	1.03	0.86	1.00	0.73	0.79	1.00	0.41	0.68	1.00	0.18	0.58	1.00
	30	0.90	0.84	1.00	0.64	0.77	1.00	0.35	0.66	1.00	0.14	0.57	1.00
1000	10	1.23	0.95	1.00	0.90	0.89	1.00	0.49	0.72	1.00	0.24	0.65	1.00
	20	1.03	0.92	1.00	0.74	0.86	1.00	0.40	0.73	1.00	0.17	0.60	1.00
	30	0.95	0.92	1.00	0.67	0.85	1.00	0.36	0.72	1.00	0.15	0.60	1.00
2000	10	1.24	0.97	1.00	0.88	0.93	1.00	0.51	0.82	1.00	0.25	0.68	1.00
	20	1.06	0.97	1.00	0.76	0.93	1.00	0.42	0.81	1.00	0.17	0.64	1.00
	30	0.95	0.96	1.00	0.67	0.92	1.00	0.37	0.79	1.00	0.15	0.64	1.00
3000	10	1.22	0.98	1.00	0.89	0.95	1.00	0.50	0.84	1.00	0.24	0.68	1.00
	20	1.05	0.99	1.00	0.76	0.96	1.00	0.42	0.85	1.00	0.17	0.67	1.00
	30	0.96	0.98	1.00	0.69	0.95	1.00	0.37	0.84	1.00	0.15	0.66	1.00

(1.75, 3.75) Çarpık Dağılım, replikasyon sayısı:100

Çizelge 20’de yer alan, iki boyutlu verilerde boyutluluk belirleme yöntemlerinden DETECT IDN indeksi sonuçları incelendiğinde, çarpık dağılım gösteren iki boyutlu veriler için, örneklem büyüklüğünün 200 ve 300 olduğu, testin uzunluğunun da 10 madde olduğu koşullarda tek boyutluluğu reddetme oranında düşüş olduğu görülmektedir. Bir başka ifadeyle testin gücünün azaldığı söylenebilir. Boyutlararası korelasyon değeri ve testin uzunluğu kaç olursa olsun, örneklem büyüklüğünün 500 ve üstü olduğu durumda tek bir koşul dışında tüm koşullarda 1.00 ret oranı elde edilmiştir. Bir başka ifadeyle her koşulda testin tekboyutlu olduğu sıfır hipotezini doğru şekilde reddetmiştir. Örneklem büyüklüğünün 500, testin uzunluğunun 10 madde, boyutlararası korelasyonun 0.90 olduğu koşul için 0.99 ret oranı sergilemiştir.

Çarpık dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu, testin uzunluğunun 10, 20 ve 30 madde olduğu ve boyutlararası korelasyonun derecesinin (0.25, 0.50, 0.75 ve 0.90) olduğu koşullarda doğrusal olmayan faktör analizi (NOHARM) istatistiği sonuçları Çizelge 21’de sunulmuştur:

Çizelge 21.

Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre Doğrusal Olmayan Faktör Analizi (NOHARM) için güç oranları

(1.75, 3.75)		Boyutlar Arası Korelasyon											
		0.25			0.50			0.75			0.90		
Örneklem Büyüküğü	Madde Sayısı	RMSR	TANAKA GFI	Reddetme oranı	RMSR	TANAKA GFI	Reddetme oranı	RMSR	TANAKA GFI	Reddetme oranı	RMSR	TANAKA GFI	Reddetme oranı
200	10	0.0048	0.9563	1.00	0.0069	0.9996	1.00	0.00452	0.9789	1.00	0.0043	0.9976	1.00
	20	0.0056	0.9821	1.00	0.0078	0.9889	1.00	0.0078	0.9789	1.00	0.0078	0.9934	1.00
	30	0.0079	0.9780	1.00	0.0088	0.9878	1.00	0.0069	0.9897	1.00	0.0056	0.9923	1.00
300	10	0.0052	0.9913	1.00	0.0065	0.9987	1.00	0.0056	0.9879	1.00	0.0115	0.9847	1.00
	20	0.0013	0.9743	1.00	0.0054	0.9791	1.00	0.0062	0.9836	1.00	0.0101	0.9893	1.00
	30	0.0023	0.9825	1.00	0.0067	0.9780	1.00	0.0097	0.9897	1.00	0.0092	0.9626	1.00
500	10	0.0063	0.9957	1.00	0.0063	0.9955	1.00	0.0052	0.9965	0.99	0.0042	0.9982	1.00
	20	0.0075	0.9887	1.00	0.0071	0.9894	1.00	0.0061	0.9926	1.00	0.0064	0.9926	1.00
	30	0.0073	0.9848	1.00	0.0074	0.9968	1.00	0.0077	0.9987	1.00	0.0059	0.9921	1.00
1000	10	0.0058	0.9841	1.00	0.0040	0.9986	1.00	0.0035	0.9986	1.00	0.0034	0.9989	1.00
	20	0.0050	0.9947	1.00	0.00647	0.9955	1.00	0.0050	0.9952	1.00	0.0052	0.9954	1.00
	30	0.0055	0.9997	1.00	0.00684	0.9993	1.00	0.0031	0.9942	1.00	0.0059	0.9966	1.00
2000	10	0.0028	0.9991	1.00	0.0028	0.9991	1.00	0.0050	0.9954	1.00	0.0027	0.9992	1.00
	20	0.0033	0.9978	1.00	0.0035	0.9973	1.00	0.0035	0.9979	1.00	0.0032	0.9983	1.00
	30	0.0056	0.9924	1.00	0.0051	0.9962	1.00	0.0059	0.9997	1.00	0.0054	0.9982	1.00
3000	10	0.0026	0.9991	1.00	0.0050	0.9954	1.00	0.0022	0.9994	1.00	0.0051	0.9955	1.00
	20	0.0030	0.9980	1.00	0.0027	0.9985	1.00	0.0029	0.9985	1.00	0.0029	0.9986	1.00
	30	0.0055	0.9996	1.00	0.0029	0.9987	1.00	0.0039	0.9997	1.00	0.0054	0.9996	1.00

(1.75, 3.75) Çarpık Dağılım, replikasyon sayısı:100

Çizelge 21'e göre çarpık dağılım gösteren verilerde testin gücü için parametrik yöntemlerden doğrusal olmayan faktör analizi (NOHARM) sonuçları incelendiğinde, çarpık dağılım gösteren iki boyutlu veriler için, boyutlararası korelasyon, örneklem büyüklüğü ve testin uzunluğu kaç olursa olsun bir koşul hariç Tanaka Uyum İyiliği İndeksi (TIGF) değeri ≥ 0.95 çıkmıştır. Bununla beraber, tüm sonuçlarda RMSR değerinin ≤ 0.05 çıkması modelin iyi uyum gösterdiğinin kanıtı olmuştur. RMSR ve Tanaka Uyum İndeksi değerleri tekboyutluluk için 1.00 ret oranı sergilemiştir. Bir başka ifadeyle her koşulda tek boyutlu olduğu sıfır hipotezini doğru şekilde reddetmiştir.

Çarpık dağılım gösteren verilerde, örneklem büyüklüğünün 200, 300, 500, 1000, 2000 ve 3000 olduğu, testin uzunluğunun 10, 20 ve 30 madde olduğu ve boyutlararası korelasyonun derecesinin (0.25, 0.50, 0.75 ve 0.90) olduğu koşullarda DIMTEST T istatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi (NOHARM) sonuçlarına için testin gücü oranları Çizelge 22'de sunulmuştur:

Çizelge 22.

Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre DIMTEST T İstatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için güç oranları

		Boyutlar Arası Korelasyon															
		0.25				0.50				0.75				0.90			
Örneklem Büyüküğü	Madde Sayısı	DIMTEST T İSTATİSTİĞİ	Boyutluluk DETECT IDN İNDEKSİ	RMSR	TIGF	DIMTEST T İSTATİSTİĞİ	Boyutluluk DETECT IDN İNDEKSİ	RMSR	TIGF	DIMTEST T İSTATİSTİĞİ	Boyutluluk DETECT IDN İNDEKSİ	RMSR	TIGF	DIMTEST T İSTATİSTİĞİ	Boyutluluk DETECT IDN İNDEKSİ	RMSR	TIGF
200	10	0.84	0.76	1.00	1.00	0.69	0.79	1.00	1.00	0.24	0.86	1.00	1.00	0.14	0.85	1.00	1.00
	20	0.96	1.00	1.00	1.00	0.89	1.00	1.00	1.00	0.36	0.99	1.00	1.00	0.08	0.99	1.00	1.00
	30	0.90	1.00	1.00	1.00	0.82	1.00	1.00	1.00	0.61	1.00	1.00	1.00	0.13	1.00	1.00	1.00
300	10	0.98	0.67	1.00	1.00	0.91	0.71	1.00	1.00	0.61	0.67	1.00	1.00	0.26	0.77	1.00	1.00
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.23	0.99	1.00	1.00
	30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.90	1.00	1.00	1.00	0.27	1.00	1.00	1.00
500	10	1.00	1.00	1.00	1.00	0.96	1.00	1.00	1.00	0.77	1.00	1.00	1.00	0.73	0.99	1.00	1.00
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.92	1.00	1.00	1.00	0.68	1.00	1.00	1.00
	30	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	0.96	1.00	1.00	1.00	0.40	1.00	1.00	1.00
1000	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.97	1.00	1.00	1.00	0.68	1.00	1.00	1.00
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.56	1.00	1.00	1.00
	30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00	0.82	1.00	1.00	1.00
2000	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00
	30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.97	1.00	1.00	1.00
3000	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00
	30	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00

Çizelge 22'ye göre, çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre DIMTEST T İstatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için güç oranları incelendiğinde, Boyutluluk DETECT IDN indeksi ve doğrusal olmayan faktör analizinin kullanıldığı tüm koşulların DIMTEST T istatistiğine göre daha doğru kararlar verdiği görülmektedir. Bununla birlikte DIMTEST T istatistiğine göre boyutlararası korelasyonun düşük olduğu koşullarda daha doğru kararlar verildiği söylenebilir. Ayrıca boyutlararası korelasyonun yüksek olduğu durumlarda DIMTEST T istatistiğinin 1000 ve üstü örneklemelerde daha iyi çalıştığı görülmektedir.

Örneklem büyüklüğünün 200 ve 300 olduğu ve bununla birlikte testin uzunluğunun 10 madde olduğu koşullarda boyutlar arası korelasyon kaç olursa olsun Boyutluluk DETECT IDN indeksinin ve DIMTEST T istatistiği sonuçlarında doğru karar verme oranının düştüğü görülmektedir. Testin uzunluğu arttıkça DETECT IDN indeksi ve DIMTEST T istatistiğinin doğru karar verme oranında artış gözlenmesine rağmen örneklem büyüklüğü azaldıkça özellikle DIMTEST T istatistiğinin doğru karar verme oranının düşük olduğu söylenebilir. Doğrusal olmayan faktör analizinin ise çarpık dağılım gösteren verilerde boyutluluk belirleme sürecinde, Boyutluluk DETECT indeksi ve DIMTEST T istatistiğine göre daha iyi çalıştığı söylenebilir.

3.3. “1-0 Puanlanan Testlerde Çeşitli Boyutluluk Belirleme Yöntemlerine Göre Standart Normal ve Çarpık Dağılım Gösteren Verilerde 1. Tip Hata Oranı Ve Testin Güç Oranları Nasıldır?”

Standart normal dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre DIMTEST T istatistiği için 1. tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerine göre DIMTEST T istatistiği için testin gücü oranı Çizelge 23'de sunulmuştur:

Çizelge 23.

Standart normal dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre DIMTEST T istatistiği için 1. tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerlerine göre DIMTEST T istatistiği için testin gücü oranları

Örneklem Büyüküğü	Madde Sayısı	I.tip hata oranı	Testin gücü oranı			
			<u>Boyutlararası korelasyon değeri</u>			
			<u>0.25</u>	<u>0.50</u>	<u>0.75</u>	<u>0.90</u>
200	10	0.04	0.91	0.64	0.37	0.14
	20	0.00	1.00	0.87	0.29	0.08
	30	0.03	1.00	0.95	0.52	0.13
300	10	0.06	0.99	0.95	0.55	0.26
	20	0.00	0.99	0.95	0.66	0.23
	30	0.00	1.00	1.00	0.86	0.27
500	10	0.02	1.00	0.97	0.69	0.30
	20	0.00	1.00	1.00	0.95	0.73
	30	0.00	1.00	1.00	0.98	0.47
1000	10	0.05	1.00	1.00	1.00	0.82
	20	0.00	1.00	1.00	1.00	0.56
	30	0.00	1.00	1.00	0.98	0.83
2000	10	0.18	1.00	1.00	1.00	0.96
	20	0.00	1.00	1.00	1.00	0.95
	30	0.01	1.00	1.00	1.00	0.98
3000	10	0.10	1.00	1.00	1.00	0.99
	20	0.00	1.00	1.00	1.00	0.98
	30	0.01	1.00	1.00	1.00	1.00

DIMTEST T istatistiği sonuçlarına göre kısa testlerde ve küçük örneklemelerde 1. tip hata oranı yüksek iken uzun testlerde ve geniş örneklemelerde 1. tip hatanın düştüğü görülmektedir. DIMTEST T istatistiğinin iki boyutlu verilerde boyutluluk kararını verirken boyutlararası korelasyondan ve kısa test uzunluğundan etkilendiği görülmektedir.

Standart normal dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Boyutluluk DETECT IDN indeksi için 1. tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerine göre Boyutluluk DETECT IDN indeksi için testin gücü oranları Çizelge 24’de sunulmuştur:

Çizelge 24.

Standart normal dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Boyutluluk DETECT IDN indeksi için 1. tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerlerine göre Boyutluluk DETECT IDN indeksi için testin gücü oranları

Örneklem Büyüklüğü	Madde Sayısı	1.tip hata oranı	Testin gücü oranı			
			<u>Boyutlararası korelasyon değeri</u>			
			<u>0.25</u>	<u>0.50</u>	<u>0.75</u>	<u>0.90</u>
200	10	0.71	0.86	0.92	0.92	0.96
	20	0.62	1.00	1.00	1.00	1.00
	30	0.53	1.00	1.00	1.00	1.00
300	10	0.64	0.74	0.85	0.91	0.97
	20	0.57	1.00	1.00	1.00	1.00
	30	0.38	1.00	1.00	1.00	1.00
500	10	0.59	1.00	1.00	0.99	1.00
	20	0.50	1.00	1.00	1.00	1.00
	30	0.38	1.00	1.00	1.00	1.00
1000	10	0.47	1.00	1.00	1.00	1.00
	20	0.43	1.00	1.00	1.00	1.00
	30	0.31	1.00	1.00	1.00	1.00
2000	10	0.39	1.00	1.00	1.00	1.00
	20	0.30	1.00	1.00	1.00	1.00
	30	0.20	1.00	1.00	1.00	1.00
3000	10	0.38	1.00	1.00	1.00	1.00
	20	0.28	1.00	1.00	1.00	1.00
	30	0.19	1.00	1.00	1.00	1.00

Çizelge 24’de yer alan DETECT IDN indeksi sonuçlarına göre, kısa testlerde ve küçük örneklemelerde 1. tip hata oranının yüksek olduğu görülürken, büyük örneklemelerde 1. tip hata oranının azaldığı söylenebilir. DETECT IDN indeksi sonuçlarına göre testin gücü oranları incelendiğinde ise DIMTEST T istatistiğinin aksine boyutlararası korelasyondan etkilenmediği söylenebilir.

Standart normal dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Doğrusal Olmayan Faktör Analizi için 1. tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerine göre doğrusal olmayan faktör analizi için testin gücü oranları Çizelge 25’de sunulmuştur:

Çizelge 25.

Standart normal dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Doğrusal Olmayan Faktör Analizi için 1. Tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerlerine göre Doğrusal Olmayan Faktör Analizi için testin gücü oranları

Örneklem Büyüküğü	Madde Sayısı	I.tip hata oranı	Testin gücü oranı			
			<u>Boyutlararası korelasyon değeri</u>			
			<u>0.25</u>	<u>0.50</u>	<u>0.75</u>	<u>0.90</u>
200	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00
300	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00
500	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00
1000	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00
2000	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00
3000	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00

Çizelge 25'e göre standart normal dağılım gösteren verilerde doğrusal olmayan faktör analizi sonuçlarının 1. tip hata ve testin gücü çalışması için diğer iki yöntemle göre en doğru kararlar verdiği ifade edilebilir.

Çarpık dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre DIMTEST T istatistiği için 1. tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerine göre DIMTEST T istatistiği için testin gücü oranları Çizelge 26'da sunulmuştur:

Çizelge 26.

Çarpık dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre DIMTEST T istatistiği için 1. Tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerlerine göre DIMTEST T istatistiği için testin gücü oranları

Örneklem Büyüklüğü	Madde Sayısı	1.tip hata oranı	Testin gücü oranı			
			<u>Boyutlararası korelasyon değeri</u>			
			<u>0.25</u>	<u>0.50</u>	<u>0.75</u>	<u>0.90</u>
200	10	0.04	0.84	0.69	0.69	0.14
	20	0.00	0.96	0.89	0.89	0.08
	30	0.00	0.90	0.82	0.82	0.13
300	10	0.05	0.98	0.91	0.91	0.26
	20	0.03	1.00	1.00	1.00	0.23
	30	0.00	1.00	1.00	1.00	0.27
500	10	0.04	1.00	0.96	0.96	0.73
	20	0.02	1.00	1.00	1.00	0.68
	30	0.01	1.00	0.99	0.99	0.40
1000	10	0.09	1.00	1.00	1.00	0.68
	20	0.00	1.00	1.00	1.00	0.56
	30	0.02	1.00	1.00	1.00	0.82
2000	10	0.16	1.00	1.00	1.00	0.99
	20	0.01	1.00	1.00	1.00	0.98
	30	0.00	1.00	1.00	1.00	0.97
3000	10	0.18	1.00	1.00	1.00	0.99
	20	0.01	1.00	1.00	1.00	0.98
	30	0.01	1.00	1.00	1.00	0.99

Çizelge 26'ya göre çarpık dağılım gösteren verilerde, DIMTEST T istatistiği için 1.tip hata oranları ve testin gücü oranları karşılaştırıldığında, 1. tip hata oranının testin uzunluğundan etkilendiğini söylemek mümkündür. Bunun yanı sıra testin gücü ise boyutlararası korelasyon değerinden etkilenmekle birlikte boyutlararası korelasyon miktarı arttıkça ve testin uzunluğu ile örneklem büyüklüğü azaldıkça testin gücünün azaldığı da görülmektedir.

çarpık dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Boyutluluk DETECT IDN indeksi için 1. tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerlerine göre Boyutluluk DETECT IDN indeksi için testin gücü oranları Çizelge 27'de sunulmuştur:

Çizelge 27.

Çarpık dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Boyutluluk DETECT IDN indeksi için 1. Tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerlerine göre Boyutluluk DETECT IDN indeksi için testin gücü oranları

Örneklem Büyüklüğü	Madde Sayısı	I.tip hata oranı	Testin gücü oranı			
			<u>Boyutlararası korelasyon değeri</u>			
			<u>0.25</u>	<u>0.50</u>	<u>0.75</u>	<u>0.90</u>
200	10	0.40	0.76	0.79	0.86	0.85
	20	0.30	1.00	1.00	0.99	0.99
	30	0.31	1.00	1.00	1.00	1.00
300	10	0.39	0.67	0.71	0.67	0.77
	20	0.21	1.00	1.00	1.00	0.99
	30	0.29	1.00	1.00	1.00	1.00
500	10	0.29	1.00	1.00	1.00	0.99
	20	0.16	1.00	1.00	1.00	1.00
	30	0.18	1.00	1.00	1.00	1.00
1000	10	0.27	1.00	1.00	1.00	1.00
	20	0.03	1.00	1.00	1.00	1.00
	30	0.08	1.00	1.00	1.00	1.00
2000	10	0.17	1.00	1.00	1.00	1.00
	20	0.01	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00
3000	10	0.06	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00

Çizelge 27'ye göre, çarpık dağılım gösteren verilerde DETECT IDN indeksi sonuçları incelendiğinde özellikle kısa testlerde 1. tip hata oranının yüksek olduğu görülmektedir. Bu hata oranı örneklem büyüklüğü arttıkça azalsa da en az hatalı sonuçların test uzunluğunun 20 ve 30 madde ve örneklem büyüklüğünün 2000 ve üstü olduğu koşulda yer aldığı söylenebilir.

Çarpık dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Doğrusal Olmayan Faktör Analizi için 1. Tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerlerine göre Boyutluluk Doğrusal Olmayan Faktör Analizi için testin gücü oranları Çizelge 28'de sunulmuştur:

Çizelge 28.

Çarpık dağılım gösteren verilerde çeşitli örneklem büyüklüğü, farklı madde sayısına göre Doğrusal Olmayan Faktör Analizi için 1. Tip hata oranı ve çeşitli örneklem büyüklüğü, farklı madde sayısı ve boyutlararası farklı korelasyon değerlerine göre Boyutluluk Doğrusal Olmayan Faktör Analizi için testin gücü oranları

Örneklem Büyükülüğü	Madde Sayısı	I.tip hata oranı	Testin gücü oranı			
			<u>Boyutlararası korelasyon değeri</u>			
			<u>0.25</u>	<u>0.50</u>	<u>0.75</u>	<u>0.90</u>
200	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00
300	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00
500	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00
1000	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00
2000	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00
3000	10	0.00	1.00	1.00	1.00	1.00
	20	0.00	1.00	1.00	1.00	1.00
	30	0.00	1.00	1.00	1.00	1.00

Çizelge 28'e göre çarpık dağılım gösteren verilerde doğrusal olmayan faktör analizi sonuçlarının 1. tip hata ve testin gücü çalışması için diğer iki yöneme göre en doğru kararlar verdiği söylemek mümkündür. 1. tip hata oranları ve testin gücünün tutarlılık gösterdiği görülmektedir.

BÖLÜM IV

4. SONUÇ ve ÖNERİLER

Çalışmanın bu bölümünde araştırmadan elde edilen sonuçlar ve sonuçlara yönelik önerilere yer verilmiştir.

4.1. Sonuçlar

1. Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T istatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için 1. tip hata oranları incelendiğinde, tekboyutluluk kararını vermede en tutarlı sonuçların Doğrusal Olmayan Faktör Analizi (NOHARM) yöntemine ait olduğu görülmektedir. Bunun yanı sıra nispeten DIMTEST T istatistiği sonuçlarının daha tutarlı olduğu görülse de kısa testlerde DIMTEST T istatistiğinin boyutluluk belirlemede kullanımının uygun olmayacağı düşünülmektedir. Bunun yanı sıra DETECT IDN indeksinin büyük örneklemelerde ve geniş madde havuzlarında kullanımının daha uygun olacağı söylenebilir. Özellikle kısa testlerde boyutluluk belirleme sürecinde DETECT IDN indeksinin kullanılmaması gerekmektedir.
2. Çarpık dağılım gösteren verilerde, çeşitli örneklem büyüklüğü ve farklı madde sayılarına göre DIMTEST T istatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi (NOHARM) için 1. tip hata oranları incelendiğinde, standart normal dağılım gösteren verilere göre çarpık dağılım gösteren verilerde DETECT IDN indeksi sonuçlarının daha tutarlı olduğu görülmektedir. DIMTEST T istatistiği ve Doğrusal Olmayan Faktör Analizi sonuçlarının ise tekboyutluluk kararını vermede daha isabetli olduğu görülmüştür.

3. Standart normal dağılım gösteren verilerde, çeşitli örneklem büyüklüğü, farklı madde sayıları ve boyutlar arası farklı korelasyon değerlerine göre DIMTEST T İstatistiği, Boyutluluk DETECT IDN indeksi ve Doğrusal Olmayan Faktör Analizi için güç oranları incelendiğinde, DIMTEST T istatistiğinin büyük örneklemli koşullarda daha doğru sonuçlar verdiği söylenebilir. Özellikle örneklem büyüklüğünün 300'den daha az olduğu durumlar için DIMTEST T istatistiğinin belirgin bir şekilde yanılma oranı artmaktadır. Aynı zamanda DIMTEST T istatistiğinin çokboyutluluk kararı için, boyutlar arası korelasyondan etkilendiği söylenebilir. Standart normal dağılım gösteren verilerde, Boyutluluk DETECT IDN indeksi ve doğrusal olmayan faktör analizi (NOHARM) sonuçlarının tüm koşullarda DIMTEST T istatistiğine göre daha doğru kararlar verdiği görülmektedir. DIMTEST T istatistiğinin ise boyutlar arası korelasyonun düşük olduğu koşullarda ve örneklem büyüklüğünün yüksek olduğu durumlarda daha doğru kararlar verdiği görülmüştür.
4. Çarpık dağılım gösteren verilerde, testin uzunluğunun 10 maddeden az olduğu koşullarda boyutluluk belirleme yöntemlerinin daha az tutarlı sonuçlar verdiği söylenebilir. Bununla birlikte DETECT IDN indeksi ve Doğrusal olmayan faktör analizi sonuçlarının kendi içinde daha tutarlı olduğu görülmekle birlikte örneklem büyüklüğünün 1000 ve üzeri olduğu koşullarda DIMTEST T istatistiğinin de boyutluluk belirlemede doğru kararlar verdiği ifade edilebilir.

4.2. Öneriler

Araştırmanın bulguları çerçevesinde benzer konularda çalışma yapacak araştırmacılara ve uygulayıcılara yönelik öneriler aşağıda sunulmuştur:

1. Bu çalışmada ele alınan koşullar; örneklem büyüklüğüne ilişkin (200, 300, 500, 1000, 2000, 3000), boyutlar arası korelasyon (0.25, 0.50, 0.75, 0.90), testin uzunluğu (10, 20, 30 madde) ve farklı yetenek dağılımları (standart normal dağılım ve çarpık dağılım) ile sınırlıdır. Benzer bir çalışma daha küçük örneklemelerde ve daha büyük test uzunluğuna sahip koşullar için tekrarlanabilir. Bunun yanı sıra araştırmaya başka değişkenler de eklenerek tekrarlanabilir.

2. Bu arařtırmada kullanılan DIMTEST T istatistięi, DETECT IDN indeksi ve doęrusal olmayan faktör analizi sonuçları temel alınarak ölkemizde uygulanan Öğrenci Seçme ve Yerleřtirme Merkezi'nin (ÖSYM) veya Milli Eğitim Bakanlığı'nın (MEB) yaptığı sınavlarda kullanılan geniş ölçekli testlerin boyutluluęunun belirlenmesinde geniş madde havuzları ve büyük örneklemeler göz önünde bulundurulduğunda doęrusal olmayan faktör analizi, DIMTEST T istatistięi ve DETECT IDN indeksinin kullanımı uygun görünmektedir. Bunun yanı sıra okul ortamında uygulanan kısa testlerin boyutluluęunun belirlenmesinde özellikle DETECT IDN indeksi ve DIMTEST T istatistięi yerine doęrusal olmayan faktör analizinin tercihi daha doęru bir karar olduęu görünmektedir.
3. Bu çalışmada 2PL ve telafisel modeller ile çalışılmıştır. Bundan sonraki çalışmalarda 3PL ve çok kategorili modeller ile birlikte özellikle bir boyutun dięer boyutu telafi etmedięi maddeleri içeren testler için, telafisel olmayan modeller kullanılarak sonuçlar incelenebilir.
4. Bu çalışmada 1-0 puanlanan test durumları oluşturulmuştur. İleride yapılacak olan çalışmalarda özellikle eğitimde ve psikolojide ölçek geliştirme ve ölçek uyarlama çalışmaları göz önünde bulundurulduğunda, çoklu puanlanan testler de aynı yöntemlerin etkililięi araştırılabilir.
5. Bu çalışmada güç çalışması iki boyutlu yapılarda ele alınmıştır. Benzer bir çalışma boyut sayısı arttırılarak da yapılabilir.
6. Aynı çalışma gerçek veriler üzerinde de yöntemlerin verimlilięi test edilebilir.
7. Bu çalışmada ele alınan testin yapısı sabitlenmiştir ve test yarı karışık (semi-mixed) yapıdadır. Basit veya karmaşık madde sayıları farklılaştırılarak daha farklı bir yapı ile benzer bir çalışma yürütülebilir ve testin yapısının etkisini sınamak için farklı test yapıları kullanılabilir.

8. Farklı madde parametre setleri, yöntemlerin performansını etkileyebilir; bu nedenle bulguların daha genellenebilir olması açısından mevcut sonuçları farklı bir madde parametresi setine dayalı sonuçlarla karşılaştırmak yararlı olabilir.
9. Ülkemizde uygulanan sınavlarda kullanılan test maddelere verilmeyen cevaplar göz önünde bulundurulduğunda, bir başka çalışmada kayıp veri miktarı manipule edilerek yöntemlerin verimliliği test edilebilir.
10. Bu çalışmada çarpık dağılım oluşturulurken Fleisman'ın (1978) çalışmasındaki çarpıklık basıklık (1.75, 3.75) değerleri dikkate alınmıştır. Standart normal dağılımdan farklı sapma miktarları dikkate alınarak veri seti oluşturulup boyutluluk belirleme süreci için 1. tip hata ve güç çalışmasına bakılabilir.
11. Bu çalışmada boyutluluk belirleme parametrik yöntemlerinden Doğrusal Olmayan Faktör Analizi (NOHARM) ve parametrik olmayan yöntemlerden DIMTEST T istatistiği ile Boyutluluk DETECT IDN İndeksi kullanılmıştır. Farklı bir çalışmada diğer parametrik ve parametrik olmayan yöntemlerin boyutluluk belirlemedeki performansı test edilebilir.
12. Bu çalışma kapsamı için seçilen parametrik ve parametrik olmayan yöntemlerden RMSR, Tanaka Uyum İyiliği İndeksi, DETECT IDN indeksi gibi indeksler kullanılmıştır. Farklı bir çalışmada aynı yöntemlerden elde edilen yaklaşık ki kare ($\chi^2_{G/D}$) istatistiği indeksi gibi diğer indekslerle de 1. tip hata ve güç çalışmasına bakılabilir.
13. Bu araştırmadan elde edilen önemli sonuçlardan biri de araştırmacılar boyutluluk belirlemede yöntemlere karar verirken veya seçim yaparken verilerin özelliği açısından yöntemlerin güçlü ve zayıf yönlerine bakmalıdır. Özellikle ülkemizde sosyal bilimler alanında veri toplama süreçlerindeki sıkıntılar göz önüne alındığında örneklem büyüklüğünün etkisinden

kaynaklanan tutarsız sonuçlara ulaşmamak için tavsiye edilen yöntemler üzerinden çalışmalar yürütülmelidir.

14. Son olarak araştırmaların henüz bir yöntemin bir diğerine üstünlüğünü kanıtlayamadığı durumlarda boyutluluk belirleme çalışması yapmak isteyen araştırmacılar, testlerden elde edilen puanlara geçmeden önce verilerin yapısını ve boyutluluğunu kapsamlı şekilde anlamak istiyorsa birden fazla boyutluluk yöntemlerinin uygulanması yararlı olabilir.





EKLER

EK 1. Güç Çalışması için Veri Üretmede Kullanılan SAS Kodları

```

data parms;
    INFILE "C:\gul_diss\parms.txt";
    INPUT a1 a2 d1;
run;
proc iml;
    call randseed(0);
    use parms;
    read all var {a1 a2} into a;
    read all var {d1} into d;

    *print a;
    *print d;

    /*generate b and theta from the standard normal distribution*/
    *b=j(30,1);
    mean = {0,0}; *change these values to transfrom thetal and theta2
    to a different mean;
    cov = {1 0.90,
           0.90 1}; *change the covariance to modify the
    correlation between dimensions;
    theta = RANDNORMAL(3000,mean,cov);

    *thetal=theta[,1];
    *theta2=theta[,2];
    * theta3=theta[,3];
    * d=(-(a1#b+a2#b));
    /*generate response data*/

    z=j(3000,30);
    p=j(3000,30);
    u=j(3000,30);

    do j=1 to 3000;
        do i=1 to 30;

            p[j,i]= exp(a[i,]*theta`[,j]+d[i,])/
            (1+(exp(a[i,]*theta`[,j]+d[i,])));

            u[j,i]=ranuni(0);
            if p[j,i]<u[j,i] then z[j,i]=0;
            else if p[j,i]>=u[j,i] then z[j,i]=1;

        end;
    end;
    create response from z;
    append from z;
    names1={thetal theta2};
    create theta from theta [colname=names1];
    append from theta;
quit;
*data keep1;
    *set bt;
    *file "C:\mirt\bpar\bpar.txt";
    *put bt;
*run;
data keep2;
    set response;
    file "C:\gul_diss\mirt_data.txt";
    put (col1-col30) (1.0);

```

```
run;
*data keep3;
*set d;
*file "C:\mirt\dpar\dpar.txt";
*put d;
*run;
proc corr data=keep2 alpha;
  var coll-col30;
  ods output cronbachalpha=alpha;
run;
data alpha_need;
  set alpha;
  if _n_=1;
  keep alpha;
  file "C:\gul_diss\alpha\alpha.txt" mod;
  put alpha;
run;
data theta;
  set theta;
  file "C:\gul_diss\theta.txt";
  put theta1 theta2;
run;
```

EK 2. 1. Tip Hata Çalışması için Veri Üretmede Kullanılan SAS Kodları

```

data parms;
    INFILE "C:\gul_diss\parms.txt";
    INPUT a1 d1;

run;
proc iml;
    call randseed(0);
    use parms;
    read all var {a1} into a;
    read all var {d1} into d;

theta=j(3000,1);
    call randgen(theta,"Normal");
z=j(3000,30);
p=j(3000,30);
u=j(3000,30);

    do j=1 to 3000;
        do i=1 to 30;
            p[j,i]= 1/(1+exp((( -1.7)*a[i,])*((theta`[,j]-
d[i,]))));

            u[j,i]=ranuni(0);
            if p[j,i]<u[j,i] then z[j,i]=0;
            else if p[j,i]>=u[j,i] then z[j,i]=1;
        end;
    end;
create response from z;
append from z;
name1={theta};
name2={d};
create theta from theta [colname=name1];
append from theta;
create d from d [colname=name2];
append from d;
quit;
*data keep1;
    *set bt;
    *file "C:\mirt\bpar\bpar.txt";
    *put bt;

*run;
data keep2;
    set response;
    file "C:\gul_diss\mirt_data.txt";
    put (col1-col30) (1.0);

run;
*data keep3;
*set d;
*file "C:\mirt\dpar\dpar.txt";
*put d;
*run;
proc corr data=keep2 alpha;
    var col1-col30;
    ods output cronbachalpha=alpha;
run;
data alpha_need;

```



```
set alpha;  
if _n_=1;  
keep alpha;  
file "C:\gul_diss\alpha\alpha.txt" mod;  
put alpha;  
run;  
data theta;  
set theta;  
file "C:\gul_diss\theta.txt";  
put theta;  
run;
```



EK 3. Testin Gücü için Yöntemlere İlişkin Programların Çalıştırılmasında Kullanılan Örnek SAS Kodları

```
options noxwait xsync;
```

```
%let nreps = 100;
```

```
%macro reps;
```

```
  %do k = 1 %to &nreps;
```

```
    %put "Starting loop &k";
```

```
    x "del C:\gul_diss\est_theta&k..txt";
```

```
    %include 'C:\gul_diss\gen_mirt_data_gul_diss.sas';
```

```
    x "copy C:\gul_diss\theta.txt C:\gul_diss\gentheta\theta&k..txt";
```

```
    x "copy C:\gul_diss\mirt_data.txt C:\gul_diss\gendata\mirt_data&k..txt";
```

```
    *x mlg_mirt.bat;
```

```
    *x "copy c:\mirt\mirt_score.sco c:\mirt\mlgtheta\mlgest_theta&k..txt";
```

```
    *x mlg_mirt_item.bat;
```

```
    *x "copy c:\mirt\mirt_item.PAR c:\mirt\mlgitem\mlgest_item&k..txt";
```

```
    x ssc.bat;
```

```
    x "copy C:\gul_diss\dimtest.out C:\gul_diss\dimtest\dimtest&k..out";
```

```
    x detect4.bat;
```

```
    x "copy C:\gul_diss\detectout.txt C:\gul_diss\detect\detectout&k..txt";
```

```
%include 'c:\gul_diss\excel_gul.sas';
```

```
%put "Ending loop &k";
```

```
  %end;
```

```
%mend;
```

```
%reps      *call macro defined above;
```

EK 4. 1. Tip Hata Çalışması için Yöntemlere İlişkin Programların Çalıştırılmasında Kullanılan Örnek SAS Kodları

```
options noxwait xsync;
```

```
%let nreps = 100;
```

```
%macro reps;
```

```
    %do k = 1 %to &nreps;
```

```
        %put "Starting loop &k";
```

```
        x "del c:\mirt\unidim\uni_est_theta&k..txt";
```

```
        %include 'c:\mirt\unidim\gen_uni_data_1.sas';
```

```
    x "copy c:\mirt\unidim\theta.txt c:\mirt\unidim\gentheta_uni\uni_theta&k..txt";
```

```
        x blm1 data_uni;
```

```
        x blm2 data_uni;
```

```
        x "copy c:\mirt\unidim\data_uni.sco  
c:\mirt\unidim\mlgtheta_uni\uni_est_theta&k..txt";
```

```
        x blm3 data_uni;
```

```
        x "copy c:\mirt\unidim\data_uni.PAR  
c:\mirt\unidim\mlgitem_uni\uni_est_item&k..txt";
```

```
        *x ssc;
```

```
        *x "copy c:\mirt\unidim\dimtestout_uni.out  
c:\mirt\unidim\dimtest_uni\uni_dimtest&k..out";
```

```
        %include 'c:\mirt\unidim\get_rmse_uni_1.sas';
```

```
%put "Ending loop &k";
```

```
%end;
```

```
%mend;
```

```
%reps          *call macro defined above;
```

KAYNAKÇA

- Ackerman, T.A. (1994). Using multidimensional item response theory to understand what items and tests are measuring. *Applied Measurement In Education*, 7(4), 255-278.
- Ackerman, T.A., Gierl, M.J., & Walker, C.M. (2003). Using multidimensional item response theory to evaluate educational and psychological tests. *Educational Measurement: Issues and Practice*, 22(3), 37-451.
- Anastasi, A. ve Urbina, S. (1997). *Psychological testing (7th ed.)*. USA: Macmillan Pub. Co. Inc.
- Cook, KF. , Kallen, MA. ve Amtmann, D. (2009). Having a fit: impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. *Applied Measurement in Education*. 18 (4): 447-460. doi: 10.1007/s11136-009-9464-4
- Crocker, L. ve Algina, J. (1986). *Introduction to classical and modern test theory*. USA: CBS College Publishing.
- Çakıcı Eser, D. (2015). *Farklı boyutluluk özelliklerindeki basit ve karmaşık yapıları testlerin çok boyutlu madde tepki kuramına dayalı parametre kestirimlerinin incelenmesi*. Yayınlanmamış doktora tezi. Hacettepe Üniversitesi, Ankara.
- Davey, T., Nering M. L. ve Thompson, T. (1997). Realistic simulation of item response data. ACT Research Report Series, 97-4. Erişim tarihi: 08.02.2014 tarihinde https://www.act.org/research/researchers/reports/pdf/ACT_RR97-04.pdf adresinden indirilmiştir.
- De Champlain, A., ve Gessaroli, M. E. (1998). Assessing the dimensionality of item response matrices with small sample sizes and short test lengths. *Applied Measurement in Education*, 11(3), 231-253.

- Embretson, S. E. ve Reise, S. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum Associates.
- Finch, H., ve Habing, B. (2005). Comparison of NOHARM and DETECT in item cluster recovery: Counting dimensions and allocating items. *Journal of Educational Measurement*, 42, 149-169.
- Finch, H., ve Habing, B. (2007). Performance of DIMTEST- and NOHARMbased statistics for testing unidimensionality. *Applied Psychological Measurement*, 31, 292-307.
- Fleisman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika*, 43, 521-532.
- Froelich, A. G., ve Habing, B. (2008). Conditional covariance-based subtest selection for DIMTEST. *Applied Psychological Measurement*, 32, 138-155.
- Froelich, A. G. (2006). A new bias correction method for the DIMTEST procedure. *Psychometrika*, 68, 435-452.
- Froelich, A. G., ve Stout, W. (2003). *A new bias correction method for the DIMTEST procedure*. Manuscript submitted for publication.
- Gessaroli, M. E., ve De Champlain, A. F. (1996). Using an approximate chi-square statistic to test the number of dimensions underlying the responses to a set of items. *Journal of Educational Measurement*, 33, 157-179.
- Göçer Şahin, S. (2016). *Çok boyutlu yapıların tek boyutlu olarak ele alınması durumunda kestirilen parametrelerin incelenmesi*. Yayınlanmamış doktora tezi. Hacettepe Üniversitesi, Ankara.
- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items*. Mahwah, nj:lawrence erlbaum.
- Hambleton, R. K., Swaminathan H., & Rogers, H.J. (1991). *Fundamentals of item response theory*. California: Sage Publications Inc.
- Hambleton, R.K., & Swaminathan, H. (1985). *Item response theory principles and applications*. Kluwer-Nijhoff Publishing. Boston-USA.
- Hattie, J. (1985). Assessing unidimensionality of tests and items. *Applied Psychological Measurement*, 9(8), 139 – 145.

- Hattie, J., Krakowski, K., Rogers, H. J., ve Swaminathan, H. (1996). An assessment of Stout's index of essential dimensionality. *Applied Psychological Measurement*, 20, 1-14.
- Hooper, D., Coughlan, J. ve Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods* 6(1), 53-60.
- Hu, L-T., ve Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: *Conventional criteria versus new alternatives*. *Structural Equation Modeling*, 6, 1-55.
- Kaya, K. Ö. ve Kelecioğlu, H. (2016). The effect of sample size on parametric and nonparametric factor analytical methods. *Kuram ve Uygulamada Eğitim Bilimleri Educational Sciences: Theory & Practice*. 16(1), 153-171.
- Kim, H. R. (1994). New techniques for dimensionality assessment of standardized test data. Unpublished doctoral dissertation. University of Illinois at Urbana-Champaign, Department of Statistics.
- Ledasma, R.D. ve Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: an easy-to use computer program for carrying out Parallel Analysis. *Practical Assessment, Research & Evaluation*, 12 (2).
- McDonald, R. P. (1981). The dimensionality of test and items. *British Journal of Mathematical and Statistical Psychology*, 34, 100-117.
- Messick, S. (1995). Validation of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meanings. *American Psychologist*, 50, 741-749.
- Mroch, A. A., ve Bolt, D. M. (2006). A simulation comparison of parametric and nonparametric dimensionality detection procedures. *Applied Measurement in Education*, 19 (1), 67-91.
- McDonald, R. P. (2000). A basis for multidimensional item response theory. *Applied Psychological Measurement*, 24, 99-114.
- McDonald, R. P. (1999). *Test Theory. A unified treatment*. Mahwah. NJ:LEA.

- Nandakumar, R. (1993). Assessing essential unidimensionality of real data. *Applied Psychological Measurement*, 17, 29-38.
- Nandakumar, R., ve Stout, W. (1993). Refinement of Stout's procedure for assessing latent trait unidimensionality. *Journal of Educational Statistics*, 18(1), 41-68.
- Özbek Baştuğ, Ö.Y. (2012). Assessment of Dimensionality in Social Science Subtest. *Educational Sciences: Theory & Practice*. 12(1), Winter: 382-385.
- Reichenberg, R.E. (2013). A Comparison of DIMTEST and Generalized Dimensionality Discrepancy Approaches to Assessing Dimensionality in Item Response Theory. M.S. dissertation, Arizona State Üniversitesi, Arizona.
- Reckase, M. D. (2009). *Multidimensional item response theory*. New York: Springer Dordrecht Heidelberg.
- Roussos, L. A., ve Özbek, O. Y. (2006). Formulation of the DETECT population parameter and evaluation of DETECT estimator bias. *Journal of Educational Measurement*, 43, 215-243.
- Seraphine, A. E. (2000). The performance of DIMTEST when latent trait and item difficulty distributions differ. *Applied Psychological Measurement*, 24, 82-94.
- Slocum-Gori, S. L. ve Zumbo, B. D. (2010). Assessing the unidimensionality of psychological scales: using multiple criteria from factor analysis. *Social Indicators Research*. 102, 443-461. Doi:10.1007/511205-010-9682-8.
- Socha, A. ve DeMars, C.E. (2013). An investigation of sample size splitting on ATFIND and DIMTEST. *Educational and Psychological Measurement*, 73, 631-647.
- Spray, J. A., Dawey, T.C., Reckase, M.D. Ackerman, T.A. & Carison, J. E. (1990). *Comparison of two logistic multidimensional item response theory models*. ACT Research report series.
- Stone, C. A. ve Yeh, C.-C. (2006). Assessing the dimensionality and factor structure of multiple-choice exams: An empirical comparison of methods using the Multistate Bar Examination. *Educational and Psychology Measurement*, 66(2), 193-214.
- Stout, W. (1987). A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika*, 52, 589-617.

- Stout, W., Habing, B., Douglas, J., Kim, H. R., Roussos, L., ve Zhang J. (1996). Conditional covariance-based nonparametric multidimensionality assessment. *Applied Psychological Measurement*, 19, 331-354.
- Stout, W., Froelich, A. G., ve Gao, F. (2001). Using resampling methods to produce an improved DIMTEST procedure. In A. Boomsma, M. A. J. van Duijn, & T. A. B. Snijders (Eds.), *Essay on item response theory* (pp. 357-375). New York: Springer.
- Sünbül, Ö. (2011). Çeşitli boyutluluk özelliklerine sahip yapılarda, madde parametrelerinin değişmezliğinin klasik test teorisi, tek boyutlu madde tepki kuramı ve çok boyutlu madde tepki kuramı çerçevesinde incelenmesi. Yayınlanmış doktora tezi, Mersin Üniversitesi Eğitim Bilimleri Enstitüsü, Mersin.
- Sünbül, Ö. ve Seo, M. (2012). *Performance of test statistics for verifying unidimensionality*, 2012 Annual Meeting, April 12-16, Vancouver, British Columbia, CANADA
- Svetina, D. (2011). Assessing dimensionality in complex data structures: A performance comparison of DETECT and NOHARM procedures. Yayınlanmış doktora tezi, Arizona State Üniversitesi, Arizona.
- Svetina, D. ve Levy, R. (2014). A Framework for Dimensionality Assessment for Multidimensional Item Response Models. *Educational Assessment*, 19(1), 35-57.
- Tate, R. (2003). A comparison of selected empirical methods for assessing the structure of responses to test items. *Applied Psychological Measurement*, 27, 159-203.
- Touron, J., Lizasoain, L. ve Joaristi, L. (2012). Assessing the unidimensionality of the School and College Ability Test (SCAT, Spanish version) using non-parametric methods based on item response theory. *High Ability Studies*. 23, 2. 183-202. DOI:10.1080/13598139.2012.735401

- Twu, B. (2005). A comparative study of DIMTEST and NOHARM in detecting the departure from unidimensionality. *Journal of Education Studies*, 39(1), 63-80.
- van Abswoude, A. A., van der Ark, L., ve Sijsma, K. (2004). A comparative study of test data dimensionality assessment procedures under nonparametric IRT models. *Applied Psychological Measurement*, 28, 3-24.
- Yen W.M. (1993). Scaling Performance Assessments: Strategies for Managing Local Item Dependence. *Journal of Educational Measurement*, Vol. 30, No. 3, pp. 187-213
- Yeh, C. C. (2007). The effect of guessing on assessing dimensionality in multiple-choice tests: A Monte Carlo study with application. Unpublished dissertation. University of Pittsburgh.
- Özer Özkan, Y. ve Acar Güvendir, M. (2014). Türkiye'de uygulanan geniş ölçekli testlerin çok boyutluluğunun analizi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 29, 31- 47.
- Yu, C. H., Popp, S.O. ve Pennell A. J. (2007). Assessing unidimensionality: a comparison of rasch modeling, paralel analysis and tetrad. *Practical Assessment, Research & Evaluation*, 12 (14).
- Yüce, G. (2012). *Seviye belirleme sınavı alttestlerinin ve ağırlıklarının önsel (a priori) ve sonsal (a posteriori) olarak belirlenmesinin bireylerin sınıflama ve sıralamaları açısından incelenmesi*. Yayınlanmamış yüksek lisans tezi. Mersin Üniversitesi, Mersin.
- Zhang, J., ve Stout, W. (1999). The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika*, 64, 213-249.
- Zhang, J. (2007). Conditional covariance theory and detect for polytomous items. *Psychometrika*, 72, 69-91.
- Zhang, B. (2008). Application of Unidimensional Item Response Models to Tests With Items Sensitive to Secondary Dimension. *The Journal of Experimental Education*, 77 (2), 147-166.

ÖZGEÇMİŞ**Kişisel Bilgiler**

Adı Soyadı : Gül GÜLER

Doğum Yeri : Arguvan

Doğum Tarihi : 15/10/1984

Eğitim Durumu

<i>Lise</i>	Malatya Anadolu Lisesi	2002
<i>Lisans</i>	İnönü Üniversitesi Sınıf Öğretmenliği	2008
<i>Yabancı Dil</i>	İngilizce: Okuma (İyi), Yazma (Orta), Konuşma	

İş Deneyimi

<i>Çalıştığı Kurumlar</i>	Gülbahçesi İlkokulu/ ADANA	2008-2008
	Mersin Üniversitesi	2008-2012
	İstanbul Aydın Üniversitesi	2012-Halen

İletişim

<i>e-Posta Adresi</i>	gulyuce2010@gmail.com@gmail.com
-----------------------	---------------------------------