T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

KISMİ EN KÜÇÜK KARELER YAPISAL EŞİTLİK MODELLERİNDE İKİLİ DEĞİŞKEN KULLANIMININ İNCELENMESİ

Mert VEZNİKLİ

DOKTORA TEZİ

İstatistik Anabilim Dalı

İstatistik Programı

Danışman

Doç. Dr. Atıf Ahmet EVREN

Temmuz, 2023

T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

KISMİ EN KÜÇÜK KARELER YAPISAL EŞİTLİK MODELLERİNDE İKİLİ DEĞİŞKEN KULLANIMININ İNCELENMESİ

Mert VEZNİKLİ tarafından hazırlanan tez çalışması 10/07/2023 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, İstatistik Programı **DOKTORA TEZİ** olarak kabul edilmiştir.

Doç. Dr. Atıf Ahmet EVREN
Yıldız Teknik Üniversitesi
Danışman

Jüri Üyeleri

Doç. Dr. Atıf Ahmet EVREN, Danışman
Yıldız Teknik Üniversitesi _____

Dr. Öğr. Üyesi Elif TUNA, Üye
Yıldız Teknik Üniversitesi _____

Dr. Öğr. Üyesi Doğan YILDIZ, Üye
Yıldız Teknik Üniversitesi _____

Dr. Öğr. Üyesi Erhan USTAOĞLU, Üye
Marmara Üniversitesi _____

Dr. Öğr. Üyesi Arzu BAYGÜL, Üye
Koç Üniversitesi _____

Danışmanım Doç. Dr. Atıf Ahmet EVREN sorumluluğunda tarafımca hazırlanan Kısmi En Küçük Kareler Yapısal Eşitlik Modellerinde İkili Değişken Kullanımının İncelenmesi başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Mert VEZNİKLİ

Her Zaman Yanımda Olan Aileme

TEŞEKKÜR

Doktora programım boyunca sağladıkları her türlü katkı için başta danışmanım **Atıf Ahmet EVREN** olmak üzere **tüm Yıldız Teknik Üniversitesi İstatistik bölümü hocalarımıza**, davetimize icabet ederek jüri ekibimizde olmayı kabul eden bölüm hocalarımız **Elif Tuna** ile **Doğan Yıldız'a**, misafir jüri hocalarımız **Erhan USTAOĞLU** ve **Arzu BAYGÜL**'e, uygulamamda verilerini kullanmama izin veren **Figen SEYMEN** ve ekibindeki **Sinem BİRANT, Yelda KASIMOĞLU ve Mine KORUYUCU'ya**, tez konumun odak noktası olan KEKK-YEM methodu ile beni tanıştıran **Behçet Yalın ÖZKARA'ya**, tez süreci boyunca verdikleri her türlü destek için **Zehra Zeynep ŞAHİNBAŞOĞLU** ve **Mehmet Şamil GÜNEŞ'e**, bu akademi yolculuğuna başlamamda bana ilk ilham kaynağı, gelişimimin her aşamasında rol model ve ışık olan **Elif Çiğdem KELEŞ'e** ve hayat boyu desteğini esirgemeyen can kardeşim **Burak KARAKUŞ'a** gönülden teşekkür ediyorum.

Mert VEZNİKLİ

İÇİNDEKİLER

TEŞEKKÜR iv
SİMGE LİSTESİ vii
KISALTMA LİSTESİ viii
ŞEKİL LİSTESİ x
TABLO LİSTESİ xi
ÖZET xii
ABSTRACT 14
1 GİRİŞ 16
1.1 Literatür Özeti ... 16
1.2 Tezin Amacı ... 19
1.3 Hipotez ... 19
2 MATERYAL VE METOD 21
2.1 Çok Değişkenli İstatistik Analiz Yöntemlerinde Temel Kavramlar ... 21
2.2 Çok Değişkenli İstatistiksel Yöntemler ... 25
2.3 Regresyon Analizi ... 28
2.4 Yapısal Eşitlik Modellemesine Giriş ... 40
2.5 Yapısal Eşitlik Modellemesine Ait Yaklaşımlar ... 43
2.6 Gizil ve Gözlemlenen Değişkenler ... 49
2.7 KEKK-YEM'de Ölçüm Modeli ... 52
2.8 KEKK-YEM'de Yapısal Model ... 57
2.9 KEKK-YEM Algoritması ... 59
2.10 KEKK-YEM Modellerinin Değerlendirilmesi ... 63
2.11 Aracılık Analizi ... 71
2.12 Düzenleyici Değişken Analizi ... 76
2.13 PROCESS Macro ... 81
2.14 KEKK-YEM Analizlerinde İncelenen Etkiler ... 82
3 UYGULAMA 87
3.1 Uygulama Verisi ... 87
3.2 İkili Bağımsız Değişken İle Nümerik Bağımlı Değişkenin Bulunduğu Model İncelemesi ... 88
3.3 İkili Bağımsız Değişken, İkili Bağımlı Değişken ve İkili Aracı Değişkenin Bulunduğu Aracılık Modeli İncelemesi ... 93

3.4 İkili Bağımsız Değişken, İkili Bağımlı Değişken ve İkili Düzenleyici Değişkenin Bulunduğu KEKK-YEM Modelinin İncelemesi..... 104

4 SONUÇ VE ÖNERİLER **113**

KAYNAKÇA **117**

TEZDEN ÜRETİLMİŞ YAYINLAR **126**

SİMGE LİSTESİ

$\boldsymbol{w}^*$	Ağırlıkların Doğrusal Birleşimi
p	Anlamlılık Değeri
a	Bileşen
Q	Blok Sayısı
$\widehat{\boldsymbol{B}}$	B'nin Tahmincisi
P	Değişken Sayısı
x	Egzojen Gizil Değişken Vektörü
β	Endojen Gizil Değişkenler Arasındaki Katsayılar Matrisi
$\boldsymbol{\xi}$	Gizil Değişken
z	Gizil Hata Terimleri Vektörü
n	Grup Frekansı
δ	Hata Terimi
ε	Hata Terimi
$\boldsymbol{\vartheta}$	İçsel Tahmin
c2	Ki-Kare Uyum Iyiliği Istatistiği
r	Korelasyon Katsayısı2.
t	T Testi Değeri
N	Toplam Örneklem Büyüklüğü
y	Vektör
f	Y Artıkları
X	Y'nin Etkili Tahmincileri
F	Y'nin Hata Matrisi
z	Z Testi Değeri

KISALTMA LİSTESİ

AVE	Açıklanan Ortalama Varyans
CR	Composite Reliability (Bileşik Güvenirlik)
CB-SEM	Covariance-Based Structural Equation Modeling (Kovaryans Tabanlı Yapısal Eşitlik Modellemesi)
CA	Cronbach's Alpha
RMSECV	Çapraz Geçerliliğe Ait Hata Kareler Ortalamasının Karekökü
MANOVA	Çok Değişkenli Varyans Analizi
d_G	Degree of Geodesic (Jeodezik Uzaklık)
d_ULS	Degree of Unweighted Least Squares (Ağırlıklandırılmamış En Küçük Kareler Derecesi)
EKK	En Küçük Kareler Yöntemi
HTMT	Heterotrait-Monotrait
KEKK-ÇGA	Kısmi En Küçük Kareler Çoklu Grup Analizi
KEKK-YEM	Kısmi En Küçük Kareler Yapısal Eşitlik Modellemesi
K-YEM	Kovaryans Tabanlı Yapısal Eşitlik Modellemesi
RMSE	Kök Ortalama Kare Hatası
LOOCV	Leave One Out Cross-Validation (Tek Bir Gözlemi Dışarıda Bırakan Çapraz Doğrulama)
MIMIC	Multiple Indicators And Multiple Causes Model (Çoklu Gösterge Ve Çoklu Neden Modeli)
NIPALS	Non-Linear Iterative Partial Least Squares (Doğrusal Olmayan Yinelemeli En Küçük Kareler)
NFI	Normed Fit Index Normlaştırılmış Uyum İndeksi
OLS	Ordinary Least Squares (En Küçük Kareler)
PLS	Partial Least Square
OR	Risk Oranı
PLS-SEM	Partial Least Square Structural Equation Modeling (Kısmi En Küçük Kareler Yapısal Eşitlik Modellemesi)
PLS-MGA	Partial Least Squares Multi-Group Analysis (Kısmi En Küçük Kareler Çoklu Grup Analizi)

PRESS	Prediction Error Sum Of Squares (Kestirim Hata Kareler Toplamı)
CR	Principal Component Regression
RR	Ridge Regresyon
SRMR	Standardized Root Mean Square Residuals (Standartlaştırılmış Hata Kareler Ortalamasının Karekökü)
SEM	Structural Equation Modeling (Yapısal Eşitlik Modellemesi)
ANOVA	Tek Yönlü Varyans Analizi
TBR	Temel Bileşenler Regresyonu
VIF	Varyans Şişirme Faktörü
YEM	Yapısal Eşitlik Modellemesi

ŞEKİL LİSTESİ

Şekil 2.1 Çok değişkenli istatistiksel yöntemlerin sınıflandırılması 26
Şekil 2.2 Lojistik fonksiyon grafikleri 39
Şekil 2.3 Formatif yapısal eşitlik modeli 54
Şekil 2.4 Reflektif yapısal eşitlik modeli 55
Şekil 2.5 MIMIC modeli 56
Şekil 2.6 Örnek yapısal model 58
Şekil 2.7 KEKK-YEM modellemesinin değerlendirme aşamaları 63
Şekil 2.8 Aracılık modelinin kavramsal diyagramı 72
Şekil 2.9 Sobel test (aracı değişken illüstrasyonu) 74
Şekil 2.10 Düzenleyici değişken modelinin kavramsal diyagramı 76
Şekil 2.11 Smartpls programında elde edilmiş bir simple slope plotun görünümü 81
Şekil 2.12 Doğrudan etkinin kavramsal diyagramı 82
Şekil 2.13 Aracılık etkisinin istatistiksel diyagramı 84
Şekil 2.14 Düzenleyicilik etkisinin istatistiksel diyagramı 86
Şekil 3.1 Birinci yapısal eşitlik modelininin smartpls gösterimi 89
Şekil 3.2 Kurulan aracılık modelinin smartpls gösterimi 93
Şekil 3.3 Model B1 için lojistik regresyon ile elde edilen b katsayılar ve standart hataları 94
Şekil 3.4 İkili değişkenlere uygulanan dönüştürme 95
Şekil 3.5 Model B2 için lojistik regresyon ile elde edilen b katsayılar ve standart hataları 98
Şekil 3.6 Model B3 için lojistik regresyon ile elde edilen b katsayılar ve standart hataları 102
Şekil 3.7 Model C için kavramsal diyagram 104
Şekil 3.8 Model C için tasarlanan KEKK-YEM'in smartpls gösterimi 105
Şekil 3.9 3 farklı modelden elde edilen p değerlerinin yönteme göre değişim grafiği 111

TABLO LİSTESİ

Tablo 3.1 Model A1 için tanımlayıcı istatistikler 90
Tablo 3.2 Model A1'dan elde edilen katsayılara ait istatistikler 90
Tablo 3.3 Model A2 için tanımlayıcı istatistikler 91
Tablo 3.4 Model A2'den elde edilen katsayılara ait istatistikler 92
Tablo 3.5 Model B1 için tanımlayıcı istatistikler 94
Tablo 3.6 Model B1'den elde edilen katsayılara ait istatistikler 96
Tablo 3.7 Model B1 için aracılık analizi sonuçları 97
Tablo 3.8 Model B2 için tanımlayıcı istatistikler 98
Tablo 3.9 Model B2'den elde edilen katsayılara ait istatistikler 99
Tablo 3.10 Model B2 için aracılık analizi sonuçları 100
Tablo 3.11 Model B3 için tanımlayıcı istatistikler 101
Tablo 3.12 Model B3'ten elde edilen katsayılara ait istatistikler 102
Tablo 3.13 Model B3 için aracılık analizi sonuçları 103
Tablo 3.14 Model C1 için tanımlayıcı istatistikler 105
Tablo 3.15 Model C1'den elde edilen katsayılara ait istatistikler 106
Tablo 3.16 Model C1 için düzenleyici değişken analizi sonuçları 106
Tablo 3.17 Model C2 için tanımlayıcı istatistikler 107
Tablo 3.18 Model C2'den elde edilen katsayılara ait istatistikler 108
Tablo 3.19 Model C2 için düzenleyici değişken analizi sonuçları 108
Tablo 3.20 Model C3 için tanımlayıcı istatistikler 109
Tablo 3.21 Model C3'ten elde edilen katsayılara ait istatistikler 110
Tablo 3.22 Model C3 için düzenleyici değişken analizi sonuçları 110

ÖZET

Kısmi En Küçük Kareler Yapısal Eşitlik Modellerinde İkili Değişken Kullanımının İncelenmesi

Mert VEZNİKLİ

İstatistik Anabilim Dalı

Doktora Tezi

Danışman: Doç. Dr. Atıf Ahmet EVREN

Bu çalışmada, normallik varsayımına ihtiyaç duymayan ve küçük örneklem boyutlarına dayanıklı olan Kısmi En Küçük Kareler Yapısal Eşitlik Modellemesi (KEKK-YEM) yönteminin, araştırmalarda sıkça kullanılan ikili değişkenler modele daha edildiğinde ne kadar tutarlı olduğunu araştırmak; doğrusal regresyon, lojistik regresyon, PROCESS Macro gibi alternatif yöntemlerle karşılaştırarak benzerliklerini, farklılıklarını, avantajlarını ve dezavantajlarını ortaya koymak ve karşılaşılan sorunlara veri dönüşümüne dayalı yeni bir öneri geliştirmek amaçlanmıştır.

Bu amaçla, gerçek veriler kullanılarak veri seti simüle edilmiş ve farklı koşullarda elde edilen sonuçlar incelenmiştir. Uygulama aşamasında, İstanbul Üniversitesi Diş Hekimliği Fakültesi Pedodonti Anabilim Dalı'nda 2016-2017 yıllarında tedavi görmüş, yaşları 3 ile 15 arasında değişen toplam 404 çocuğa ait veriler üzerinde çalışılmıştır. Bu verilerde, çocukların cinsiyeti, biberon ve emzik kullanımı, kötü ağız alışkanlıkları ve bazı ortodontik rahatsızlıkların varlığı gibi kategorik değişkenler bulunmaktadır.

Yapılan analizler sonucunda KEKK-YEM ile klasik doğrusal regresyonun aynı ve/veya çok yakın sonuçlar ürettiği, fakat ikili değişken grupları arasında frekans

farkı çok yüksek olduğunda KEKK-YEM'in doğrusal regresyona kıyasla tip 2 (beta tipi) hataya daha dirençli olduğu görülmüştür.

KEKK-YEM ve Kısmi En Küçük Kareler Çoklu Grup Analizi (KEKK-ÇGA) yöntemlerinin, dummy kodlama ile kullanıldıklarında, grup başına düşen örneklem büyüklerinin eşit veya birbirlerine yakın olduğu durumlarda, Sobel testleri ve PROCESS Macro gibi alternatif yöntemlerle benzer sonuçlar verdiği ve aracılık ile düzenleyicilik etkilerini tespit edebildiği görülmüştür. Diğer taraftan grup başına düşen örneklem büyüklerinin 30'un altında kaldığı ve gruplar arasındaki farkın çok yüksek olduğu (7-8 kat kadar) durumlarda aracılık modelleri için KEKK-YEM, düzenleyicilik modelleri için ise KEKK-ÇGA yöntemi tutarsız sonuçlar verebilmektedir. Bu tip durumlarda değişkenlere dummy kodlama yerine nümerik kodlama dönüşümü uygulamak daha güvenilir ve daha güçlü sonuçlar elde etmeye yardımcı olabilir.

Anahtar Kelimeler: kısmi en küçük kareler yapısal eşitlik modellemesi, kısmi en küçük kareler çoklu grup analizi, lojistik regresyon, process macro, ikili değişken

ABSTRACT

Investigation of the Use of Binary Variables in Partial Least Squares Structural Equation Models

Mert VEZNİKLİ

Department of Statistics

Doctor of Philosophy Thesis

Supervisor: Assoc. Prof. Atıf Ahmet EVREN

This study aims to investigate the consistency of Partial Least Squares Structural Equation Modeling (PLS-SEM) which does not rely on the assumption of normality and is robust to small sample sizes, when applied to binary variables commonly used in research. Furthermore, it aims to compare PLS-SEM with alternative methods such as linear regression, logistic regression, and PROCESS Macro to uncover their similarities, differences, advantages, and disadvantages, and to develop a novel proposal based on data transformation to address encountered issues.

To achieve these objectives, real data was used, and a simulated dataset was examined under different conditions. In the implementation phase, the data of a total of 404 children, aged between 3 and 15, who received treatment in the Istanbul University Faculty of Dentistry in the years 2016-2017, were analyzed. The dataset included categorical variables such as the gender of the children, bottle and pacifier usage, presence of poor oral habits, and certain orthodontic disorders.

The analysis results indicated that PLS-SEM and classical linear regression produced similar and/or very close results, but when there was a high frequency difference between binary variable groups, PLS-SEM demonstrated higher statistical power than linear regression. When PLS-SEM and Partial Least Squares Multi-Group Analysis (PLS-MGA) methods were used together with dummy coding, it was observed that they provided similar results to Sobel tests and alternative methods such as PROCESS Macro when the sample sizes per group were equal or close to each other. They were also able to identify mediating and moderating effects. However, in cases where the sample sizes per group were below 30 and there was a significant difference between the groups (around 7-8 times difference), PLS-SEM could yield inconsistent results for mediation models, while PLS-MGA could yield inconsistent results for moderation models. In such cases, applying numerical coding transformation to the variables instead of dummy coding may provide more reliable and robust results.

Keywords: partial least squares structural equation modeling, partial least square multi group analyze, logistic regression, process macro, binary variable

1
GİRİŞ

1.1 Literatür Özeti

Çok değişkenli istatistiksel yöntemler, modern istatistik ve veri bilimi alanlarında çok önemli bir rol oynar. Bu yöntemler, birçok değişken arasındaki ilişkileri inceleyerek, bu ilişkileri anlama ve analiz etme konusunda bize yardımcı olurlar. Bu nedenle, çok değişkenli istatistiksel yöntemler, araştırmacıların karşılaştığı veri analizi sorunlarını çözmek için yaygın olarak kullanılır.

Gelişen teknolojiler ve büyük veri setlerinin ortaya çıkması, çok değişkenli istatistiksel yöntemlerin kullanımını daha da önemli hale getirmiştir. Bu yöntemler, büyük veri setlerindeki çok sayıda değişken arasındaki karmaşık ilişkileri daha etkili bir şekilde analiz etmek için kullanılabilir.

Son yıllarda, makine öğrenmesi ve yapay zekâ tekniklerinin gelişmesiyle birlikte, çok değişkenli istatistiksel yöntemlerin önemi daha da artmıştır. Bu teknikler, karmaşık veri setlerindeki çok sayıda değişkeni bir arada değerlendirerek, tahmin yapma ve sınıflandırma gibi birçok farklı amaç için kullanılabilir.

Yapısal eşitlik modellemesi de çok değişkenli istatistiksel bir analiz yöntemidir ve birçok alanda kullanılmaktadır. YEM, bir dizi ölçülebilir değişken arasındaki ilişkileri modellemeye ve analiz etmeye izin verir. Bu yöntem, birçok disiplinde, özellikle sosyal bilimler, psikoloji, yönetim ve pazarlama gibi alanlarda kullanılmaktadır. Bilim, teknoloji, mühendislik ve matematik eğitimi[1], sosyal bilimler[2], sağlık[3], çevrimiçi alışverişte tüketici davranışı[4] ve mikro, küçük ve orta ölçekli işletmelerin performansı ile ilgili çalışmalarda uygulanmıştır. YEM, önemli bir ampirik kanıt temeli ve altta yatan teorinin güçlü bir kavramsal anlayışı olduğunda karmaşık, bütünleşmiş teorik modelleri test etmek için özellikle yararlıdır.

YEM, 2 adet temel bileşen içerir: yapısal model ve ölçme modeli. Yapısal model, birbirleriyle ilişkili değişkenler arasındaki doğrudan ve dolaylı ilişkileri

açıklamaktadır. Ölçme modeli ise, değişkenlerin ölçülmesine ilişkin yapısal modeldeki ilişkileri yansıtır ve ölçümlerdeki hata unsurlarını hesaba katar.

YEM, doğrudan ve dolaylı etkileri modelleme kabiliyeti ile diğer yöntemlerden farklıdır. Örneğin, regresyon analizi, yalnızca doğrudan etkileri modelleyebilirken, YEM, dolaylı etkileri de modelleyebilir. Ayrıca birçok değişkeni bir arada değerlendirmeye olanak tanır ve bu nedenle karmaşık ilişkileri analiz etmek için de kullanılabilir. Bir araştırmacı, bir hipotezi test etmek veya bir kuramı doğrulamak için YEM kullanılabileceği gibi, aynı zamanda bir ölçek veya testin geçerliliği ve güvenirliği gibi özelliklerini değerlendirmek için de kullanılabilir.

YEM, kesikli veya sürekli verilerle kullanılabilir. Ayrıca, yapısal denklem modelleri ve faktör analizi gibi diğer istatistiksel tekniklerle birleştirilebilir.

YEM için tercih edilebilen iki yöntem vardır. Bunlar en çok olabilirlik yöntemi kullanılan Kovaryans Tabanlı Yapısal Eşitlik Modellemesi (Covariance-Based Structural Equation Modeling) yaklaşımı ("K-YEM"–"CB-SEM") ve Varyans Tabanlı Kısmi En Küçük Kareler Yapısal Eşitlik Modellemesi (Partial Least Squares Structural Equation Modeling) yaklaşımı ("KEKK-YEM"-"PLS-SEM") dir. K-YEM, veri normal dağılıma uyduğunda etkilidir ve parametre keşfi için maksimum olabilirlik yöntemlerini kullanır. KEKK-YEM ise veriler normal dağılmadığında bile etkilidir ve değişkenler arasındaki ilişkilerin açıklanmasına daha fazla odaklanır. KEKK-YEM, model özelliklerini tahmin etmek için bileşen tabanlı bir yaklaşım kullanırken, K-YEM, varsayımlara uygunluğu değerlendirmek için genellikle çoklu uyum indeksleri kullanır. K-YEM, değişkenlerin ölçek türü açısından daha esnektir ve hem nominal hem de sayısal değişkenlerle çalışabilir. KEKK-YEM ise, sayısal değişkenlerin ölçeği açısından daha esnektir, ancak nominal değişkenlerle çalışmak daha zordur. K-YEM, modeldeki parametrelerin karmaşıklığına daha iyi uyum sağlar ve özellikle büyük ve karmaşık yapısal modellerde daha iyi performans gösterirken, KEKK-YEM, modeldeki değişkenler arasındaki ilişkiyi daha iyi yakalar ve küçük örneklemlerde daha iyi performans gösterir.

K-YEM daha geleneksel bir yaklaşım olarak kabul edilirken, KEKK-YEM daha yenilikçi bir yaklaşımdır. KEKK-YEM'in avantajları arasında, yukarıdaki bahsedilen varsayımların bu yöntem için sıkı olmaması, gözlemlenen değişkenlerin sayısının

gizil değişkenlere göre daha yüksek olabileceği, çoklu bağımsız değişkenlerin modelde kullanılabileceği, çoklu ölçütlerin kullanılabileceği, normal olmayan dağılımların işlenebileceği ve hipotez testleri için yeniden örnekleme (bootstrap) tekniklerinin kullanılabileceği, hata terimlerinin normal olmama, değişen varyans ve otokorelasyona duyarsız olması[5] sayılabilir. Bu sebeplerden ötürü, çok sayıda disiplinde kullanılabilir ve giderek daha popüler hale gelmiştir. Özellikle işletme yönetimi, pazarlama, finans, insan kaynakları yönetimi gibi disiplinlerde yaygın olarak kullanılmaktadır. Bilgi sistemleri, e-ticaret, bilgi yönetimi, bilgi güvenliği ve diğer ilgili alanlarda kullanılabilir. Endüstri mühendisliği, yazılım mühendisliği ve diğer mühendislik alanlarında kullanılabilir. Sağlık bilimleri, sağlık yönetimi, tıp ve diğer ilgili alanlarda kullanılabilir. Sosyal psikoloji, eğitim, sosyoloji, politika bilimi ve diğer sosyal bilimlerde kullanılabilir. Sürdürülebilirlik, ekoloji ve diğer çevre bilimleri alanlarında kullanılabilir. Tarım ekonomisi, tarım politikası, gıda bilimi ve diğer tarım bilimleri alanlarında kullanılabilir.

YEM modelleri, günümüzün tıbbi istatistiğinde de ilgili regresyon denklemleri sistemini çözmek için giderek daha fazla kullanılmaktadır. Bununla birlikte, daha geniş kullanımı için büyük bir engel, kategorik değişkenleri genelleştirilmiş doğrusal modeller çerçevesinde ele almanın zorluğu olmuştur. Birkaç araştırmacı, KEKK-YEM'in genellikle nominal, sıra, aralık ve oran ölçekli değişkenlerle çalıştığını vurgulamaktadır[6-8]. Ancak bu uygulama dikkatle değerlendirilmelidir. Örneğin, araştırmacılar bir seçim durumunu belirtmek için bir endojen yapıyı ölçmek için ikili tek bir gösterge kullanmaya karar verebilirler. Ancak bu durumda, gizli yapı bağlı olduğu değişkenin ölçüsü haline gelir ki bu da yol katsayıları Sıradan En Küçük Kareler Yöntemi regresyonları ile tahmin edildiğinden KEKK-YEM algoritmasındaki yaklaşımlar için sorunludur[9].

Lojistik regresyon, tıbbi istatistiklerde kategorik sonuçları modellemenin temel taşı haline gelmiş olsa da sonuçların doğal olarak ilişkili olduğu durumlarda bile pragmatik bir yaklaşım olarak pek sorgulanmamaktadır. Bu durum, aynı değişkenin, hem bir ya da birkaç değişkenin etkisi altındaki bağımlı değişken, hem de başka bir ya da birkaç değişkenin tahmincisi olabileceği süreç değerlendirmesinde de yaygındır. Örneğin, erken doğum, hem hamilelik sürecinin önemli bir sonucu hem de düşük doğum ağırlığı sebebi ile gelecekteki sağlığı

olumsuz etkileyebilen bir risk faktörü olabilir. Bu sonuçların sıralı doğası, aynı sonucu farklı zaman noktalarında ele alan tekrarlanan ölçüm modelleri tarafından kapsanmaz. Lojistik regresyon modeliyle ele alınması zor olan bir araştırma problemine başka bir örnek, bir sonucun yalnızca öngörücü değişkenlerin doğrudan etkileriyle değil, aynı zamanda 2 değişken arasındaki ilişkiyi düzenleyen düzenleyici değişkenler tarafından belirlendiği zamanlardır. Bazı mide ilaçlarının iyileştirici etkisinin, hastaların ilacı aldıkları andaki açlık tokluk durumuna göre ortaya çıkması buna güzel bir örnektir. Bir deneyde ikili bir sonucun olasılığını açıklamaya çalışan bazı araştırmacılar, kimi çalışmalarda bir bağımsız değişkeninin etkisinin ikinci bir değişken tarafından ılımlı hale getirilmesiyle ilgilenirler. Bu tür verileri analiz etmek için en yaygın olarak kullanılan yöntem lojistik regresyondur, ancak bu yöntem bağımlı değişkeni doğrusal olmayan bir dönüşüme tabi tuttuğundan, ortaya çıkan etkileşim katsayıları orijinal olasılıklardaki ılımlılık etkilerini düzgün bir şekilde yansıtmaz. Anlamlı düzenleyici etkileri, anlamlı olmayan etkileşim katsayıları ile sonuçlanabilir ve bu tersi durum için de geçerlidir.

1.2 Tezin Amacı

Bu çalışmanın amacı, bir önceki başlıkta bahsedilmiş olan, ikili değişkenlerin YEM modellerin kullanımı ile ilgili zorluğun ana yönlerini gözden geçirmek, normallik varsayımı gerektirmeyen ve çok küçük örneklemlere dahi dirençli olan KEKK-YEM'in tıbbi istatistikte sıkça kullanılan ikili değişkenlerle kullanılan modellerde ne kadar tutarlı olduğunu araştırmak ve doğrusal regresyon, lojistik regresyon, PROCESS Macro gibi alternatif yöntemlere göre benzerliklerini, farklılıklarını, avantaj ve dezavantajlarını ortaya koymak ve bu soruna veri dönüşümüne dayalı yeni bir öneri geliştirmektir. Bunun için gerçek veriler kullanılacak olup, veri seti simüle edilerek farklı şartlarda nasıl sonuçlar elde edildiği incelenmiştir.

1.3 Hipotez

Bilimsel araştırmalarda sıklıkla kullanılan değişkenlerden biri, örneğin cinsiyet gibi iki uçlu nominal değişkenlerdir. Diğer taraftan, çalışmalarda tasarlanan araştırma modelleri günümüzde giderek daha fazla kompleks hale gelmekte ve ilişkiler örgüsünü analiz etmek için parametrik olmayan ve küçük örneklemlerle

güçlü istatistik sonuçları verebilen yapısal eşitlik modeli yaklaşımı KEKK-YEM'in kullanımı revaçtadır. Başta tıp alanında olmak üzere laboratuvar ortamında gerçekleştirilen bilimsel çalışmaların çoğunda gerek etik gerek maddi sebeplerle örneklem büyüklüğünün kısıtlayıcı bir faktör olması da bilinen bir gerçektir. Bu nedenle KEKK-YEM modellerinde iki uçlu nominal değişkenlerin (ikili değişkenlerin) doğru şekilde kullanılabilmesi, hem araştırmacılar için bir seçenek olabilmesi hem de hem istatistiki çalışmaların daha güvenilir olabilmesi adına önemlidir. Çalışmada grup başına düşen örneklem büyüklüğü çarpık olmaması şartıyla hem küçük hem büyük örneklemler için ikili değişkenlerin, bağımlı, bağımsız, aracı ve düzenleyici değişken olarak dummy kodlama ile kullanılabileceği savunulmaktadır. Gruplar arasında çarpık dağılım olduğunda ise dummy kodlama yerine normal dağılımdan faydalanarak yapılacak bir dönüşüm ile nümerik kodlama uygulanmasının daha doğru bir yaklaşım olabileceği iddia edilmektedir.

2
MATERYAL VE METOD

2.1 Çok Değişkenli İstatistik Analiz Yöntemlerinde Temel Kavramlar

2.1.1 Değişken

Değişkenleri anlamak, çeşitli araştırma alanlarında çalışmaları tasarlama, verileri analiz etme ve sonuçları yorumlama açısından temel öneme sahiptir. İstatistikte, değişkenler birkaç türe ayrılabilir.

2.1.1.1 Kategorik Değişken

Kategorik değişkenler, özelliklerine göre kategorilere ayrılabilen değişkenlerdir. Ayrıca niteliksel ya da nominal değişkenler olarak da bilinirler. Cinsiyet, okul türü ve konaklama türleri gibi örnekler verilebilir[10].

Kategorik değişkenlerin, grup sayısına bağlı ikili değişken ve kukla değişken gibi özel halleri de mevcuttur.

- İkili (binary) değişken, sadece iki farklı kategori veya seviyeye sahip olan bir değişkendir. Bu kategoriler genellikle "var" veya "yok", "evet" veya "hayır", "başarılı" veya "başarısız" gibi zıt kutuplu kavramları ifade eder. İkili değişkenler özellikle araştırmalarda sık kullanılan bir türdür ve genellikle kategorik değişkenler arasında yer alırlar. İkili değişkenler sıklıkla ve birçok alanın (tıbbi araştırmalar, sosyal bilimler, pazarlama, ekonomi ve diğer) analizlerinde kullanılır. Örneğin, bir araştırmacı bir tedavi yönteminin etkililiğini ölçmek için bir ikili değişken olan "iyileşmeyi” kullanabilir ve sonuçları analiz etmek için uygun istatistiksel yöntemleri uygulayabilir.
- Kukla (gölge) değişken, bir kategorik değişkenin 2 veya daha fazla kategorik grubu arasındaki etkisini modellemek için kullanılan bir tür sahte (dummy) değişkendir. Bu, bir bağımsız değişkenin kategorik gruplarını temsil eden 1 veya daha fazla dummy değişken oluşturarak yapılır. Kukla değişkenler, özellikle regresyon analizinde kullanılır ve bir kategorik bağımsız değişkenin etkisini kontrol etmek için kullanılabilirler. Kukla değişkenler, çaprazlamaları

sıfıra eşitlemek için de kullanılabilir ve bu nedenle birçok farklı analiz türü için yararlıdır. Bu tür değişkenlerin analizi ve yorumlanması, istatistiksel analizlerde farklı yöntemler gerektirebilir. Örneğin, ikili değişkenlerin ortalamaları hesaplanamaz ve sıklıkla yüzdelik oranları kullanılır.

Bilinmelidir ki ikili değişken ve kukla değişken kavramları benzer olsalar dahi birbirleri ile karıştırılmamalıdırlar. İkili değişkenler, sadece 2 farklı kategorik değere sahip olan değişkenlerdir. Örneğin, bir öğrencinin cinsiyeti (erkek/kadın) veya bir yargıcın kararı (suçlu/suçsuz) ikili değişken olabilir. Bu değişkenler 0/1, 1/2 gibi farklı şekillerde isteğe göre kodlanabilir.

Diğer taraftan kukla değişkenler, 1 veya daha fazla kategorik değişkenin tek bir veya birkaç alt kümesini temsil eden 0 ve 1 değerlerini alan yapay değişkenlerdir. Kukla değişkenler 2, 3 veya çok daha fazla sayıda kategoriye sahip olabilir ve genellikle bir kategorik değişkenin n farklı kategorisi varsa, n-1 tane kukla değişken oluşturulur. Bu değişkenler, asıl kategorik değişkenin tüm kategorilerini temsil etmezler ancak diğer tüm kategorilerin bir kombinasyonunu temsil ederler. Bu nedenle, bir kukla değişkeni oluşturulduğunda, diğer kategoriler varsayılan olarak 0 olarak kabul edilir. Kukla değişkenler, birçok çok değişkenli analiz yönteminde kullanılan önemli bir araçtır çünkü kategorik değişkenleri niceliksel analizlerde kullanılabilecek formatlara dönüştürmelerine ve kategorik değişkenlerin etkilerini nicel olarak hesaplayabilmelerine olanak tanırlar.

2.1.1.2 Nümerik Değişken

Nümerik değişkenler, sayısal değerlere sahip olan ve nicel bilgileri ifade eden değişkenlerdir. Bu tür değişkenler, ölçülebilir bir özelliği temsil ederler ve kesirli veya tam sayı değerleri alabilirler. Bir kişinin yaşı, boyu, ağırlığı veya geliri örnek olarak verilebilir. Bu tür değişkenler sayılarla ifade edilir ve matematiksel işlemlere tabi tutulabilirler. Nümerik değişkenler genellikle sürekli veya kesikli olabilirler.

- Sürekli değişkenler, belirtilen bir aralıkta herhangi bir değeri alabilen değişkenlerdir. Nicel ya da aralık değişkenleri olarak da bilinirler. Örnekler arasında bir ailenin aylık geliri, bir tedaviye harcanan tutar verilebilir[10].

- Kesikli değişkenler, genellikle tam sayılar şeklinde belirli değerler alabilen değişkenlerdir. Nicel bir değişken türüdürler. Örnek olarak bir ailedeki birey sayısı ve bir hastaneye yapılan ziyaret sayısı verilebilir[10].

Değişkenler bir ilişki model içerisinde sahip oldukları role göre sınıflandırıldığında ise şu şekilde gruplandırılabilirler.

2.1.1.3 Bağımsız Değişken

Bu değişkenler, bağımlı değişken üzerindeki etkilerini belirlemek için bir çalışmada manipüle edilen ya da kontrol edilen değişkenlerdir. Ayrıca tahminci ya da açıklayıcı değişkenler olarak da bilinirler[11].

2.1.1.4 Bağımlı Değişken

Bağımlı değişkenler, bir çalışmada ölçülen sonuçlardır. Bağımsız değişkenler tarafından etkilenirler ve yanıt ya da sonuç değişkenleri olarak da adlandırılırlar[11].

2.1.1.5 Kukla Değişken

Gizil (gizli) değişkenler, doğrudan gözlemlenemeyen veya ölçülemeyen, ancak diğer gözlemlenebilir değişkenlerden çıkarılan değişkenlerdir[12].

2.1.1.6 Gözlemlenen Değişken

Bir çalışmada doğrudan ölçülebilen ya da gözlemlenebilen değişkenlerdir. Ayrıca belirgin ya da ölçülmüş değişkenler olarak da bilinirler[12].

YEM gibi çok değişkenli analizlerde, çoklu bağımsız ve bağımlı değişkenler aynı anda analiz edilerek gizli değişkenler ve aralarındaki etkiler ölçülebilir.

2.1.2 Kodlama

Değişken kodlama, istatistiksel analizlerde kategorik değişkenlerin sayısal değerlere dönüştürülmesi işlemidir. Bu işlem, kategorik değişkenlerin analizlerde kullanılabilmesini sağlar. Kodlama yöntemi, kullanılan analiz yöntemine ve veri setine bağlı olarak değişebilir.

Kodlama yöntemleri arasında en yaygın olanları, nominal kodlama, ordinal kodlama, dummy kodlama ve "effect coding" olarak sayılabilir.

İstatistikte nominal kodlama, kategorik verileri sayısal formatta temsil etmek için kullanılan bir yöntemdir. Bu kodlama yöntemi, her kategoriye bir sayı atayarak gerçekleştirilir. Bu sayılar arasında matematiksel işlemler yapılamaz, çünkü sayılar yalnızca kategorileri temsil etmek için kullanılır. Örneğin, bir araştırmada cinsiyet değişkeni nominal kodlama ile temsil edilebilir, erkekler için 1 ve kadınlar için 2 gibi. Ancak bu kodlamada, 1 ve 2 sayıları yalnızca cinsiyet kategorilerini temsil etmek için kullanılabilir ve bu sayılar arasında matematiksel işlemler yapılamaz.

Ordinal kodlama, sıralı veya dereceli değişkenlerin sayısal değerlere dönüştürülmesini ifade eder. Bu kodlama yöntemi, sıralı değişkenlerin analiz edilmesinde kullanılır ve değişkenin sıralı doğası korunur. Örneğin, "çok az", "az", "orta", "çok" ve "çok fazla" gibi sıralı değerleri olan bir değişkeni kodlamak için, her kategoriye bir sayı atanır. Örneğin, "çok az" 1, "az" 2, "orta" 3, "çok" 4 ve "çok fazla" 5 gibi. Bu sayılar ardışık ve eşit aralıklı değildir, ancak birbirinden farklı ve sıralıdır.

Ordinal kodlama, değişkenler arasındaki ilişkileri daha doğru bir şekilde yansıtmaya yardımcı olurken, değişkenin oran ölçeğine sahip olmadığına dikkat etmek önemlidir. Bu nedenle, oranlar veya oranlar üzerinde yorum yapmak için kullanılmamalıdır.

Dummy kodlama, çok değişkenli istatistik analizlerinde kategorik değişkenlerin analiz edilebilmesi için kullanılan önemli bir yöntemdir. Çok değişkenli regresyon analizleri gibi yöntemlerde, kategorik değişkenlerin analiz edilebilmesi için bu değişkenlerin sayısal bir forma dönüştürülmesi gerekmektedir. Bu dönüşüm için dummy kodlama sıklıkla kullanılır. Bu yöntem, değişkenin her bir kategorisi için ayrı bir ikili (0-1) değişken oluşturarak gerçekleştirilir. İlk adımda, orijinal kategorik değişkenin her bir kategorisi için ayrı bir ikili değişken oluşturulur. Örneğin, bir "cinsiyet" değişkeni için "erkek" ve "kadın" kategorileri varsa, iki ayrı ikili değişken oluşturulur: "erkek" için bir ikili değişken oluşturulur ve bu değişkenin değeri "1" olurken, "kadın" için ayrı bir ikili değişken oluşturulur ve bu değişkenin değeri "1" olur. "Erkek" için oluşturulan ikili değişkenin "kadın" için oluşturulan ikili değişkenden farkı, "erkek" ikili değişkenin değerinin "1" olduğu

durumlarda "kadın" ikili değişkeninin değerinin "0" olmasıdır ve bu tam tersi içinde geçerlidir. Bu şekilde, her bir kategorik değişken kategorisi için ayrı bir ikili değişken oluşturularak, orijinal kategorik değişken sayısal bir formata dönüştürülür ve her bir gözlem, kategorik değişkenin bir kategorisi için 1 ve diğer kategoriler için 0 olan bir dizi dummy değişkeni ile temsil edilir. Bu dizi, regresyon modeline dahil edilerek kategorik değişkenin etkisi incelenebilir. Dummy kodlama, birçok alanda kullanılan bir yöntemdir. Örneğin, pazarlama araştırmalarında tüketicilerin demografik özelliklerinin analizinde, tıp alanında hastalıkların risk faktörlerinin incelenmesinde ve sosyal bilimlerde sosyo-ekonomik faktörlerin etkilerinin araştırılmasında yaygın olarak kullanılır. Bu nedenle, dummy kodlama, çok değişkenli istatistik analizlerinde kategorik değişkenlerin analiz edilmesi için önemli bir araçtır.

2.2 Çok Değişkenli İstatistiksel Yöntemler

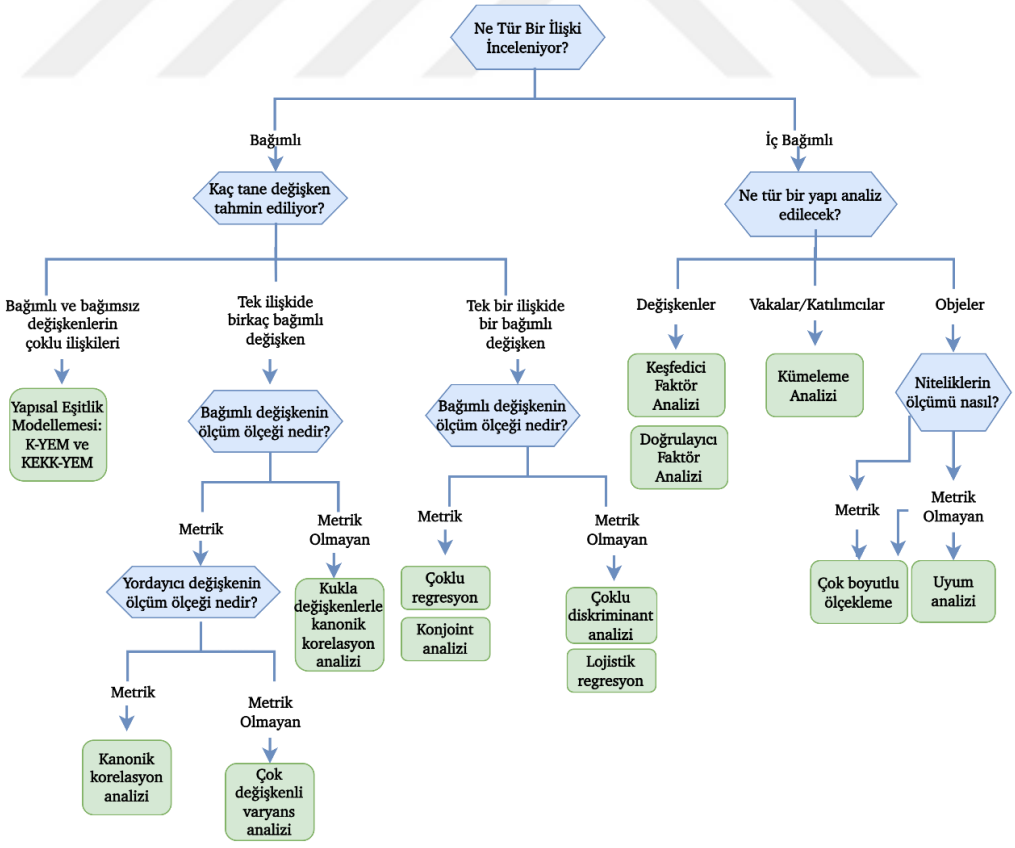
Birçok araştırma disiplininde, aynı anda birden fazla değişkenin ölçüldüğü ve analiz edildiği durumlarla karşılaşmak yaygındır. Geleneksel istatistiksel yöntemler, tek değişkenli analizleri ele almak için tasarlanmış olup birden fazla değişken arasındaki karmaşık ilişkileri ve etkileşimleri yakalamada yetersiz kalabilir. İşte bu noktada, çok değişkenli istatistiksel yöntemler devreye girer.

Çok değişkenli istatistiksel yöntemler, araştırmacılara birden fazla değişken arasındaki ilişkileri aynı anda keşfetme ve anlama gücü sağlar. Bu yöntemler, desenleri, bağımlılıkları ve değişkenler arasındaki etkileşimleri inceleyerek karmaşık olguların daha kapsamlı analizine imkân tanır. Değişkenlerin birlikte davranışını dikkate alarak, çok değişkenli yöntemler, verinin temel yapısını ve dinamiklerini anlama konusunda değerli bir içgörü sunar.

Çok değişkenli istatistiksel yöntemlerin temel hedeflerinden biri, gözlenen değişkenler arasındaki ilişkilerin temel nedenlerini veya gizli yapıları belirlemek ve analiz etmektir. Bu gizli yapılar, doğrudan ölçülemeyen kavramları veya boyutları temsil eder ve ölçülen değişkenlerden çıkarılabilir. Bu gizli yapıları yakalayarak, araştırmacılar gözlenen verileri etkileyen temel mekanizma ve süreçleri daha iyi anlayabilirler.

Çok değişkenli istatistiksel yöntemlerde modeller “bağımlı” ve “iç bağımlı” olmak üzere 2 ayrı gruba ayrılır. Bu ayrımı yapmak için sorulacak temel sorulardan biri “Değişkenler, bazı teorilere dayanarak bağımsız ve bağımlı olarak sınıflandırılabilir mi?” sorusudur. Bağımlı modellerde, hangi değişkenlerin bağımsız hangilerinin bağımlı olduğu açık bir şekilde belli olduğu ve bir değişkenin veya değişken setinin, bağımsız değişkenler olarak bilinen diğer değişkenler tarafından tahmin edilmesi veya açıklanması amaçlanır. Örneğin çoklu regresyon analizi, bir “bağımlılık tekniği” örneğidir. Buna karşılık, iç bağımlılık tekniği, tek bir değişkenin veya değişken grubunun bağımsız veya bağımlı olarak tanımlanmadığı bir tekniktir. Bu yöntem, tüm değişkenlerin aynı anda analiz edilmesini içerir. Keşfedici faktör analizi, iç bağımlılık tekniği örneğidir[13].

Yöntem seçiminde sorulacak diğer 2 soru ise “tek bir analizde kaç bağımlı değişken olduğu” ve “değişkenler nasıl ölçüldüğü” sorularıdır. Çok değişkenli yöntemlerin bu ayrımlara göre nasıl sınıflandırıldığı Şekil 2.1’de aşamalandırılmıştır[13].



Şekil 2.1 Çok değişkenli istatistiksel yöntemlerin sınıflandırılması

Çok değişkenli yöntemler, geniş bir teknik yelpazesini kapsar ve her birinin kendine özgü güçlü yönleri ve uygulama alanları bulunur. Yaygın olarak kullanılan bazı çok değişkenli yöntemler arasında faktör analizi, küme analizi, diskriminant analiz, çok değişkenli regresyon ve yapısal eşitlik modellemesi yer alır. Bu teknikler, çok değişkenli verilerde desenleri ortaya çıkarma, varyansı açıklama, sonuçları tahmin etme ve hipotez testi gibi farklı yaklaşımlar sunar.

En çok kullanılan yöntemlerden biri olan Çok Değişkenli Varyans Analizi (MANOVA), ANOVA'dan farklı olarak tek bir tane değil, iki ya da fazla bağımlı değişkenin aynı anda analizine izin verir. Bağımsız değişken düzeylerinin, kendi başlarına ve/veya kombinasyon halinde bağımlı değişkenleri etkileyip etkilemediğini belirlemek için kullanılır[14].

Çoklu Doğrusal Regresyon, bağımlı değişken ile bağımsız değişkenler arasındaki ilişkinin doğrusal olduğu çoklu regresyonun bir türüdür. Birden fazla bağımsız değişken içeren doğrusal denklemin katsayılarını tahmin eder ve bağımlı değişkenin değerini en iyi tahmin eden denklemi sağlar[14].

Lojistik Regresyon, ikili bağımlı değişken ile bir veya daha fazla bağımsız değişken arasındaki ilişkiyi analiz etmek için kullanılan bir istatistik yöntemidir. Bağımsız değişkenlerin değerlerine dayanarak bir olayın olasılığını tahmin eder [14].

Diskriminant analizi, değişkenlerin değerlerine dayanarak gözlemleri iki veya daha fazla gruba sınıflandırmak için kullanılan çok değişkenli bir istatistiksel tekniktir. Hangi değişkenlerin grupları ayırt ettiğini belirlemeye yardımcı olur, sınıflandırma ve tahmin için kullanılabilir[14].

Faktör analizi, birçok değişkeni daha az sayıda temel faktöre indirmek için kullanılan bir istatistik yöntemidir. Gözlemlenen değişkenler arasındaki ilişkileri açıklayan gizli değişkenleri belirler ve veri yapısını basitleştirmeye yardımcı olur[14].

Küme analizi, birden fazla değişken üzerindeki benzerliklerine göre gözlemleri gruplandırmak için kullanılan bir tekniktir. Veri içindeki desenleri ve ilişkileri belirlemeye yardımcı olur ve bölümlendirme, sınıflandırma ve veri azaltma için kullanılabilir[14].

Çalışmanın odak noktası olan yapısal eşitlik modellemesi ise Çok değişkenli verilerin analizi ve gizli değişkenlerin ve bunların arasındaki ilişkilerin ölçülmesi için kullanılan bir istatistiksel teknikler ailesidir. Faktör analizi ve çoklu regresyon analizini birleştirerek, karmaşık ve bütünsel teorik modellerin test edilmesine imkan sağlar. Literatürde Yol (path) Analizi olarak da ifade edilmektedir. Yol analizi ilk olarak 1930'lu yıllarda genetikçi Sewall Wright tarafından geliştirilen bir yaklaşım olup, YEM modellerindeki değişkenler arasındaki istatistiksel ilişkilerin ayrıştırılmasında kullanılmaktadır [12].

Çok değişkenli istatistiksel yöntemlerin uygulama alanları çeşitlidir ve birçok disiplinde yaygın olarak kullanılır. Sosyal bilimlerde çok değişkenli yöntemler, anket verilerini analiz etmek, karmaşık sosyal olguları incelemek ve tutumlar, davranışlar ve sosyo-ekonomik faktörler gibi birden fazla değişken arasındaki ilişkileri anlamak için kullanılır. İşletme ve pazarlama araştırmalarında çok değişkenli teknikler, tüketici tercihlerini ortaya çıkarmak, pazarları segmente etmek ve öngörü modelleri geliştirmek için kullanılır. Mühendislik ve çevre bilimlerinde çok değişkenli yöntemler, karmaşık sistemleri analiz etmeye, değişkenleri izlemeye ve çeşitli faktörler arasında bilinçli kararlar almaya yardımcı olur.

Bu çalışmanın odak noktası olan yapısal eşitlik modellemesi temelde bir regresyon modellemesi olduğu için ve çalışmanın uygulama kısmında farklı türde regresyon modellemeleri de kullanıldığı için, Bölüm 2.3'te bu konu ayrıntılı bir biçimde açıklanmıştır.

2.3 Regresyon Analizi

İstatistiğin bir araştırma aracı olarak kullanıldığı gerek sosyal gerek fen bilimlerinden birçok bilim dalında, araştırmacılar birbirleri ile ilişki içinde olan olguları tespit etmek ve tespit ettikleri bağımlı ile bağımsız değişkenler arasındaki ilişkiyi incelemek isterler. Bu doğrultuda regresyon modelleri günümüzde en çok tercih edilen ve faydalanılan analiz çeşitlerinden biridir. Regresyon analizinin birçok çeşidi bulunmaktadır. Bunlardan bazıları aşağıda sıralanmıştır[15].

- Doğrusal Regresyon
- Poisson Regresyon

- Negatif Binom Regresyon
- Lojistik Regresyon
- Probit Regresyon
- Temel Bileşenler Regresyonu
- Ridge Regresyon
- Cox Regresyon

Sebep-sonuç ilişkileri üzerine odaklı yapılan çalışmalarda, eldeki veriye ve amaca uygun olan doğru istatistiki yöntemi kullanmak son derece önemlidir. Yukarıda listelenen regresyon modellerinden **doğrusal regresyon ve lojistik regresyon** en sık karşılaşılan regresyon türlerindendir.

2.3.1 Doğrusal Regresyon

Modelin bağımlı değişkeni sürekli bir değişken ise genellikle doğrusal regresyon kullanılır.

Klasik bir doğrusal regresyon modeli şu şekilde gösterilir.

$$\boldsymbol{y} = \boldsymbol{\beta_0} + \boldsymbol{X\beta} \ + \boldsymbol{\varepsilon} \tag{2.1}$$

Bu gösterim, X matrisine ait değişkenlerin ve y vektörüne ait değişkenlerin ortalamaları çıkarılarak merkezîleştirildiğinde, eş. 2.2'de olduğu gibi basitleştirilebilir.

$$\boldsymbol{y} = \boldsymbol{X\beta} \ + \boldsymbol{\varepsilon} \tag{2.2}$$

Birden fazla bağımlı değişken olduğundaysa modelin matris gösterimi şu şekildedir.

$$\boldsymbol{Y} = \boldsymbol{X\beta} \ + \boldsymbol{E} \tag{2.3}$$

Doğrusal regresyonda B'yi kestirmek amacıyla sıklıkla tercih edilen metot, En Küçük Kareler yöntemi/EKK (Ordinary Least Squares/OLS) idir. Eğer X matrisi p rankına sahip tam ranklı bir matrisse, β'nın EKK tahmincisi $\hat{B}_{\text{EKK}}$, şu şekilde hesaplanır.

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{EKK}} = (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\,\boldsymbol{X}'\boldsymbol{y} \tag{2.4}$$

$\hat{B}$, β'nın elemanlarının yansız kestirimlerini verir. İki ya daha fazla bağımlı değişken olduğunda EKK tahmin edicisi,

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{EKK}} = (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\,\boldsymbol{X}'\boldsymbol{Y} \tag{2.5}$$

şekilde genelleştirilebilir[16].

Doğrusal modellerde bir dizi varsayım bulunmaktadır. Bu varsayımlar, hata terimlerinin normal olarak dağıldığı, hata terimlerinin beklenen değerinin ortalamasının sıfır olduğu ve varyansının homojen olduğu (σ2'ye eşit), hataların bağımsız olduğu, hata terimleri ile bağımsız değişkenler arasında korelasyon olmadığı gibi sıralanabilir.[17].

Modelde birden çok bağımsız değişken olduğunda bu varsayımlara ilave olarak, açıklayıcı değişkenlerin birbirinden bağımsız olması varsayımının da sağlanması gerekmektedir[18]. Diğer bir ifade ile çoklu bağlantı durumunun olmaması gerekir. Bu varsayım eş. 2.5'de görülebileceği üzere X'X matrisinin tersinin alınması ile ilgilidir.

Regresyon katsayılarını EKK yöntemiyle tahmin etmek için, p×p boyutlu X'X matrisinin tersinin alınabilmesi gerekmektedir. Bu matrisin tersi alınabilmesi için, X değişkenlerinin doğrusal olarak bağımsız olması gibi belirli koşulların sağlanması gerekmektedir. Aynı zamanda, p değişken sayısının n gözlem sayısına eşit ya da daha az olması da başka bir önemli koşuldur ($p \leq n$).

Eğer değişkenler arasında çoklu bağlantı varsa, X'X matrisinin tersi alınamaz ve EKK tahmincisinin değişkenliği yüksek çıkar. Bunun sonucunda, yüksek bir R^2 değerine rağmen, tahmin edilen EKK katsayıları istatistiksel olarak anlamsız olabilir. Bu katsayılar çok büyük, çok küçük veya yanlış işarete sahip olabilir. Bu yüzden, Kısmi En Küçük Kareler/Partial Least Squares (KEKK/PLS), Ridge Regresyon (RR) ve Temel Bileşenler Regresyonu/Principal Component Regression (TBR/CR) gibi alternatif yöntemler, artan yanlılık ve azalan varyans arasında bir denge sağlamak amacıyla kullanılır[16, 19].

KEKK, analiz veya tahmin amacıyla bir bağımlı değişken kümesini bağımsız değişken kümesinden tahmin etmek için kullanılan bir yöntemdir. Bu tahmini elde etmek için, en iyi tahmin gücüne sahip olan dik bileşenler kümesi kullanılır ki bu bileşenler aynı zamanda gizil değişkenler olarak da adlandırılır[20]. KEKK, çoklu bağlantı sorununu çözmek için tercih edilen yanlı regresyon yöntemlerinden biridir. TBR ve RR gibi yöntemler, çoklu bağlantı sorunuyla başa çıkmak için orijinal bağımsız değişken sayısından daha az faktör kullanır.

2.3.2 Kısmi En Küçük Kareler Regresyonu

EKK, bağımsız değişkenlerin fazla sayıda ve birbiri ile ilişkili olduğu durumlarda yeni bir tahmin yöntemi olarak geliştirilmiştir[21]. KEKK yöntemi, X ve Y blokları arasındaki gizli değişken bağlantısından faydalanarak veri blokları arasındaki ilişkiyi bulmayı amaçlar. Başlangıçta ekonomik ve sosyal olayları modellemek için kullanılan KEKK, daha sonra Svante Wold'un çalışmalarıyla kimya bilimi alanında popüler hale gelmiştir[22].

EKK yöntemi, bağımlı ve bağımsız değişkenler arasındaki ilişkiyi kurmak için yeni bağımsız değişkenler olan faktörler, gizli değişkenler veya bileşenler oluşturur. Her bir bileşen, bağımsız değişkenlerin lineer bir kombinasyonudur. Daha sonra, standart regresyon yöntemleri kullanılarak bu bileşenleri bağımlı değişkenlere bağlayan denklemler belirlenir. KEKK yöntemi, temel bileşenler regresyonundan farklı olarak, bileşenlerin oluşturulmasında X ve Y değişkenlerine ait veri değerlerinin etkisi olduğu bir yöntemdir. KEKK yöntemi, regresyon probleminin boyutunu azaltmak için bağımsız değişkenlerin sayısından daha az sayıda bileşen kullanırken, aynı zamanda bağımlı değişkenleri tahmin etmek için X değişkenlerindeki bilginin büyük bir kısmını yakalamayı amaçlar[23].

2.3.2.1 Kısmi En Küçük Kareler Regresyon Modeli

Regresyon analizinde, bağımsız değişken sayısının azaltılması önemli bir görevdir. Yeni değişkenler ekledikçe, regresyon katsayılarının varyansları genellikle düşmez. Bu nedenle, değişkenleri çıkarmak, uygulamada bir avantajdır. KEKK, regresyon analizi için iyi bir yöntem olarak kabul edilebilir, çünkü bağımlı değişkenleri kesin bir şekilde açıklayan bileşenleri seçer. KEKK regresyon modeli tarafından bulunan gizli değişkenler, X skorları olarak da adlandırılır ve t_a

(a=1,2,...,A) ile gösterilir. X skorları, Y'nin tahmin edicileridir ve aynı zamanda X'i modellemektedir. EKK yönteminde t_a değerleri, Doğrusal Olmayan Yinelemeli En Küçük Kareler (NIPALS/Non-linear Iterative Partial Least Squares) algoritması kullanılarak tekrarlayan bir şekilde belirlenebilir. KEKK yönteminin en dikkate değer özelliği, değişkenler yüksek bir derecede bağlantılı veya doğrusal olarak bağımsız olsalar bile az bileşenli basit bir model bulabilmesidir[24].

X skorları, birbirleriyle dik olacak şekilde sınırlı sayıda bileşenden oluşur. Bu skorlar, orijinal x_k değişkenleri ile $w_{ka}^{*}(a = 1,2, \dots, A)$ katsayıların (ağırlıkların) doğrusal birleşimi olarak kestirilir. Bu katsayılar zaman zaman r_{ka} ile ifade edilebilir. X skorları şu şekilde hesaplanır.

(Vektör biçiminde)

$$\mathbf{t_{ia}} = \sum_{\mathbf{k}} \mathbf{w_{ka}^{*}} \mathbf{x_{ik}} \tag{2.6}$$

(Matris biçiminde)

$$T=XW^{*} \tag{2.7}$$

X skorlarının özellikleri şunlardır.

- X skorları, X'in etkili bir şekilde özetlendiği p_{ak} yükleriyle çarpılarak hesaplanır. Bu durumda, X artıkları minimize edilir.

(Vektör biçiminde)

$$\mathbf{x_{ik}} = \sum_{\mathbf{a}} \mathbf{t_{ia}} \mathbf{p_{ak}} + \mathbf{e_{ik}} \tag{2.8}$$

(Matris biçiminde)

$$X = TP' + E \tag{2.9}$$

Çok değişkenli Y için (M>1) u_a olarak ifade edilen Y skorları, Y'nin başarılı bir şekilde özetlendiği c_{am} ağırlıklarıyla çarpılır. Böylece, g_{im} artıkları en aza indirgenir.

(Vektör biçiminde)

$$\mathbf{y_{im}} = \sum_{\mathbf{a}} \mathbf{u_{ia}} \mathbf{c_{am}} + \mathbf{g_{im}} \quad (2.10)$$

(Matris biçiminde)

$$Y=UC' + G \quad (2.11)$$

- X skorları, Y'nin etkili bir şekilde tahmincileridir. Y artıkları olarak adlandırılan f_{im} değerleri, gözlenen ve tahmin edilen bağımlı değişkenler arasındaki farklılıkları simgeler ve Y'nin hata matrisi olan F'nin elemanlarını oluşturur.

(Vektör biçiminde)

$$\mathbf{y_{im}} = \sum_{\mathbf{a}} \mathbf{c_{ma}} \mathbf{t_{ia}} + \mathbf{f_{im}} \quad (2.12)$$

(Matris biçiminde)

$$Y=TC'+F \quad (2.13)$$

Bu denklemler, eş. 2.6 ve eş. 2.7 kullanılarak, çoklu regresyon modeline benzer bir şekilde şu şekilde dönüştürülebilir.

(Vektör biçiminde)

$$\mathbf{y_{im}} = \sum_{\mathbf{a}} \mathbf{c_{ma}} \sum_{\mathbf{k}} \mathbf{w^{*}_{ka}} \mathbf{x_{ik}} + \mathbf{f_{im}} = \sum_{\mathbf{k}} \mathbf{b_{mk}} \mathbf{x_{ik}} + \mathbf{f_{im}} \quad (2.14)$$

(Matris biçiminde)

$$Y = XW^{*}C'+F = XB+F \quad (2.15)$$

Regresyon katsayıları ise şu şekilde hesaplanır.

(Vektör biçiminde)

$$\mathbf{b_{mk}} = \sum_{\mathbf{a}} \mathbf{c_{ma}} \mathbf{w^{*}_{ka}} \tag{2.16}$$

(Matris biçiminde)

$$B = W^{*}C' \tag{2.17}$$

Tek bir y değişkeni (M=1) ve X'X köşegen bir matris olduğu durumda, X'deki değişkenler arasında ilişki yapısı bulunmamaktadır. KEKK regresyon, tek bir bileşende Çoklu Lineer Regresyon çözümüne ulaşır. Bu şekilde, iki yöntemin de katsayıları, $w_1c'_1$ olarak eşitlenir.

Her bileşen (a) sonrasında, X matrisinin $t_ap'_a$ yı çıkarılarak indirgenir. Bu sayede KEKK modelinin w_a ağırlıkları şeklinde ifade edilmesi sağlanır. Bu işlemde X değişkenleri yerine önceki boyuttaki E_{a-1} hata terimleri kullanılır. Bu sayede şu eşitlikler yazılabilir.

(Vektör biçiminde)

$$\mathbf{t_{ia}} = \sum_{\mathbf{k}} \mathbf{w_{ka}} \mathbf{e_{ik,a-1}} \tag{2.18}$$

(Matris biçiminde)

$$\mathbf{t_a} = \mathbf{E_{a-1}} \mathbf{w_a} \tag{2.19}$$

Y matrisi de $t_ac'_a$ yı çıkarılarak indirgenebilir, ancak bu indirgeme zorunlu değildir. Sonuçlar, Y'nin indirgenip indirgenmediğine bağlı olmaksızın benzerdir.

(Vektör biçiminde)

$$\mathbf{e_{ik,a-1}} = \mathbf{e_{ik,a-2}} - \mathbf{t_{i,a-1}} \mathbf{p_{a-1,k}} \tag{2.20}$$

$$\mathbf{e_{ik,0}} = \mathbf{x_{ik}} \tag{2.21}$$

(Matris biçiminde)

$$\mathbf{E_{a-1} = E_{a-2} - t_{a-1} p'_{a-1}} \quad (2.22)$$

$$\mathbf{E_0 = X} \quad (2.23)$$

w ağırlıkları, w* olarak gösterilen ağırlıklara dönüştürülebilir. ∗w ağırlıkları X ile doğrudan ilişkilidir ve Eş. 2.6'yı sağlar. Bu iki ağırlık arasında $W^*=W(P'W)^{-1}$ şeklinde bir ilişki bulunur.

2.3.2.2 Doğrusal Olmayan Yinelemeli En Küçük Kareler Algoritması

1984 yılında Wold ve arkadaşlarının geliştirdiği [25], KEKK'nın temel algoritması olan Doğrusal Olmayan Yinelemeli En Küçük Kareler/Non-linear Iterative Partial Least Squares (kısa gösterimi ile NIPALS) algoritması, isteğe bağlı olarak dönüştürülmüş, ölçeklendirilmiş ve merkezleştirilmiş orijinal X ve Y verileriyle başlar. Algoritmanın adımları şu şekilde listelenebilir.

1. Sıklıkla, başlangıç vektörü olarak Y matrisinin sütunlarından biri kullanılır ve u ile temsil edilir. Eğer yalnızca tek bir y değişkeni varsa, u=y olarak belirlenir.
2. X ağırlıkları w: w=X'u/u'u olarak ifade edilir. Burada w esnek bir şekilde değiştirilebilir. w ağırlıklarına, w=w/||w|| şeklinde normalizasyon uygulanır.
3. X skorları olan t hesap edilir. T=Xw
4. Y ağırlıkları, c: c=Y't/t't olarak hesaplanır ve c değerlerine de c=c/||c|| şeklinde normalizasyon uygulanır.
5. Y skorlarının güncellenmiş bir kümesi, u: u=Y'c/c'c hesaplanır.
6. t'deki değişimi kullanarak yakınsaklık testi yapılır. Örneğin, $||t_{eski}- t_{yeni}||/||t_{yeni}|| < \varepsilon$ şeklinde kontrol edilebilir. Burada ε değerinin, 10^{-6} ile 10^{-8} aralığında küçük bir değer olması gerekir. Eğer yakınsaklık sağlanmazsa ikinci adıma dönülür, sağlanırsa 7. Aşamaya geçilir ve daha sonra tekrar 1. Aşama ile devam edilir. Eğer M=1 olduğunda sadece tek bir y değişkeni varsa c değeri 1'e eşit olur, süreç tek bir yinelemede yakınsar. Daha sonra ise doğrudan 7. aşamaya geçilir.
7. Mevcut bileşenler X ve Y'den çıkarılır. Bu indirgenmiş matrisler, bir sonraki bileşende X ve Y olarak kullanılır. Y'nin indirgenmesi isteğe bağlıdır ve sonuçlar Y'nin indirgenip indirgenmediğine bağlı olmaksızın benzerlerdir.

X yükleri: p=X't/(t't)

Y yükleri: q=Y'u/(u'u)

Regresyon (t üzerine u'nun): b=u't/(t't)

Hata matrisleri: X→X-tp' ve Y→Y-btc'

8. Çapraz doğrulama ile X'de Y hakkında daha fazla anlamlı bilgi olmadığını gösterilene kadar, bir sonraki bileşenle algoritma devam ettirilir (ilk aşamaya geri dönülür).

Gözlemlenen artık matrisleri olan yeni X ve Y matrisleriyle bir sonraki yineleme başlar. Yinelemeler, bir durdurma ölçütü kullanılana kadar veya X sıfır matrisi haline gelene kadar devam edebilir.

Modelin veriye uyumunun ve kestirici gücünün belirlenmesinde, farklı ölçümler literatürde yer almaktadır. RMSE (kök ortalama kare hatası), modelin veriye uyumunu gösteren bir metriktir ve şu şekilde hesaplanır.

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\boldsymbol{y_i} - \hat{\boldsymbol{y}}_i)^2}{\boldsymbol{n}}} \qquad (2.24)$$

Eşitlikteki $\hat{y}_i$ değerleri, tüm gözlemler model kurumunda yer aldığında kestirilen değişkenden elde edilen değerlerdir. Çapraz geçerliliğe ait hata kareler ortalamasının karekökü (RMSECV/root mean square error of cross-validation), modelin yeni gözlemleri tahmin etme kabiliyetini gösterir. RMSECV, RMSE ile benzer şekilde tanımlanır. Ancak RMSECV'de, $\hat{y}_i$ değerleri model oluşumunda yer almayan gözlemler için tahminlerdir. RMSECV, modelde bulunan gizil değişken sayısı için kestirim hata kareler toplamına (PRESS/prediction error sum of squares) bağlıdır ve

$$\boldsymbol{RMSECV} = \sqrt{\frac{\boldsymbol{PRESS}}{\boldsymbol{n}}} \qquad (2.25)$$

şeklinde hesaplanır

PRESS, modelin tahmin hatalarının kareler toplamını ifade eden bir istatistiksel ölçüdür. Genellikle, PRESS değeri ne kadar düşükse modelin tahmin yeteneği o kadar yüksektir. Bu değerini kestirmek için genellikle tek bir gözlemi dışarıda

bırakan çapraz doğrulama (LOOCV/leave one out cross-validation) yöntemi tercih edilir. Çapraz doğrulama, bileşenleri seçmek için modelin tahmin yeteneğini maksimize etmek amacıyla kullanılan bir yöntemdir. LOOCV, her seferinde bir gözlemi modele dahil etmeyerek ve bir kez tahmin ederek çapraz doğrulama işlemini gerçekleştirir.

Non-linear Iterative Partial Least Squares (NIPALS), Kısmi En Küçük Kareler Yapısal Eşitlik Modellemesi (PLS-SEM) içinde gizil değişkenlerin tahmininde sıklıkla kullanılan bir yöntemdir. KEKK-YEM, kısmi en küçük kareler regresyonu ve yapısal eşitlik modellemesi kavramlarını birleştiren istatistiksel bir modelleme tekniğidir.

KEKK-YEM'de, NIPALS algoritması, modelin gizil değişkenlerini ya da bileşenlerini yinelenen bir şekilde tahmin etmek için kullanılır. Bu gizil değişkenler, modelde doğrudan gözlemlenemeyen ancak gözlemlenen göstergelerden çıkarılan altta yatan yapılardır.

NIPALS algoritması, KEKK-YEM'de kovaryansı maksimize eden bileşenleri iteratif olarak çıkarır. İlk bileşen için ağırlıklar tahmin edilir, ardından bu bileşen için skorlar hesaplanır. Bu skorlar, gözlemlenen değişkenleri azaltmak için kullanılır ve süreç istenen bileşen sayısına kadar tekrarlanarak ardışık bileşenler elde edilir.

Elde edilen bileşenler daha sonra yapısal eşitlik modelinde yol katsayılarını ve model parametrelerini tahmin etmek için kullanılır. Bu yol katsayıları, gizil değişkenler ile gözlemlenen göstergeler arasındaki ilişkileri temsil eder. KEKK-YEM'de tahmin süreci, gözlemlenen göstergelerin açıklanan varyansını ve modelin verilere uyumunu maksimize etmeyi amaçlar.

2.3.3 Lojistik Regresyon

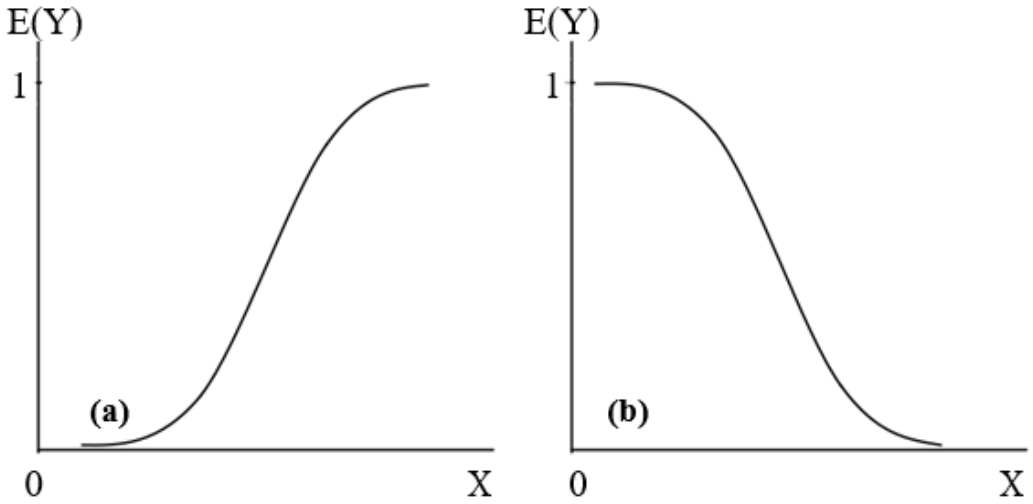
Doğrusal regresyon modelinde bağımsız değişkenlerin sürekli ya da kategorik olması, kullanılacak yöntemler ve elde edilecek parametre tahminleri üzerinde etkili değildir. Ancak diğer taraftan, bağımlı değişkenin kategorik bir yapıda olması tüm model yapısını değiştirmektedir. Çünkü modelin bağımlı değişkeni kategorik olduğunda doğrusal regresyonun varsayımlarından olan normallik

ortadan kalkmakta, En Küçük Kareler Yöntemi kullanılamamakta, bu nedenle doğrusal regresyon kullanabilme şansı ortadan kalkmaktadır[26].

Bağımlı değişken, ikili veya daha çok sayıda gruplu bir yapıya sahip olduğunda kullanılabilecek çok sayıda regresyon çeşidi vardır. Doğrusal regresyondan en büyük farkı, bağımlı değişkenin kategorik yapıda olması olan lojistik regresyon da bu türlerden biridir ve sıklıkla karşımıza çıkmaktadır. Doğrusal ve lojistik regresyon türü arasındaki bu farklılık, hem parametrik model seçimini hem de varsayımları etkilemektedir[17]. Doğrusal regresyon modeli bağımlı değişkenin alacağı değer tahminlenirken, lojistik regresyon, bağımlı değişkenin alabileceği herhangi bir değerin gerçekleşme olasılık yüzdesi tahminlenmektedir. Bu nedenle lojistik regresyonda temel amaç, bağımlı değişkenin alabileceği değerlere dair olasılığı tahminleyen bir tahmin modeli kurmak ve bu model vasıtasıyla bir duruma dair olasılığı gösteren ve dolayısıyla 0 ila 1 arasında değişen bir olasılık değeri hesaplayabilmektir. Diğer bir deyişle model bu değerlerinde bir risk hesaplayamaz. Ancak bu durumunun her şartta her model için geçerli olmadığını da belirtmek gerekir[26].

Lojistik Regresyon Analizi, bir regresyon modeli türetmekle birlikte aynı zamanda bir ayırıcı analiz tekniği olma özelliği de taşır. Buna bağlı olarak, lojistik regresyonu bir regresyon analizlerden farklılaştıran 3 önemli özellik bulunmaktadır[27].

- Regresyon Analizinde bağımlı değişken sayısal, Lojistik Regresyon Analizinde nitel bir değişkendir.
- Regresyon Analizinde bağımlı değişkenin değeri, Lojistik Regresyon ’da ise bağımlı değişkenin alabileceği değerlerden birinin gerçekleşme olasılığı kestirilir.
- Regresyon Analizinde bağımsız değişkenlerin çoklu normal dağılım göstermesi koşulu var iken, Lojistik Regresyon ’un uygulanabilmesi için bağımsız değişkenlerin dağılımına ilişkin hiçbir koşul bulunmamaktadır.



Şekil 2.2 Lojistik fonksiyon grafikleri

Lojistik regresyon fonksiyonu,

$$\boldsymbol{\pi(x) = \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)}} \tag{2.26}$$

ya da

$$\boldsymbol{\pi(x) = [1 + \exp(\beta_0 + \beta_1 X)]^{-1}} \tag{2.27}$$

olarak yazılabilir[27].

$\pi(x) = E(Y/x)$ değeri şartlı ortalama olarak bilinir.

Şartlı ortalamanın, modelde yer alan parametrelerle ($\beta o + \beta 1$) doğrusal hale dönüştürülmesi için, transformasyona tabi tutulması gerekir. Bu transformasyona Logit transformasyon adı verilir ve aşağıdaki şekilde gösterilir:

$$\boldsymbol{\pi(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)}\right] = \beta_0 + \beta_1 X} \tag{2.28}$$

Transformasyon değişkeni g(x), modeldeki parametrelerle doğrusaldır, süreklidir ve $-\infty$, $+\infty$ aralığında değişen değerler alır. $\pi(x)$ arttıkça g(x)'te artar ve eğer $\pi(x) < 0.5$ ise g(x) negatif, $\pi(x) > 0.5$ ise g(x) pozitif değerler alır[17].

Modelin sonuç değişkeninin sınırlarını genişletmek için uygulanan Logit transformasyonunun bazı özellikleri şöyle sıralanabilir:

- p arttıkça lojit(p) de artar.

- p, 0 ile 1 arasında iken lojit(p) reel sayılar doğrusu üzerinde değerler alabilir.

- $p < 0.5$ olduğunda lojit(p) < 0 ve
$p > 0.5$ olduğunda lojit(p) > 0 olur[28].

2.4 Yapısal Eşitlik Modellemesine Giriş

Yapısal eşitlik modellemesinin tarihsel gelişimi, regresyon analizinin oluşturulmasına kadar uzanmaktadır.

İlk modeller, doğrusal regresyon modellerini içerir ve bir regresyon ağırlığını hesaplamak için korelasyon katsayısı ve en küçük kareler kriterini kullanır. Karl Pearson, 1896 yılında iki değişken arasındaki ilişkiyi belirtmede kullanılan korelasyon katsayısı için bir endeks sağlayan bir formül geliştirmiştir. Regresyon modeli, bağımlı değişken puanlarının (Y puanları) tahmin edilebilmesi için bağımsız gözlenen puanların (X puanlarının) doğrusal bir şekilde ağırlıklandırılmasıyla karesel artık hata miktarını en aza indirir[29].

Bundan sadece 8 yıl sonra Charles Spearman, Pearson'ın 1904'te keşfettiği korelasyon katsayısını, öğelerin birbiriyle ilişkilerini gruplandırmak ve tanımlamak için kullanmaya başlamıştır. Bu yaklaşımda, öğelerin arasındaki ilişkiler ve birlikte değişimler göz önüne alınarak, doğrulayıcı faktör analizi kavramı ortaya atılmış ve 1940'larda bir istatistiksel ölçüm aracı olarak geliştirilmiştir[30].

Faktör analizi geliştirmesinde de yer alan Sewall Wright aynı zamanda yol (path) analizinin de geliştiricisidir. Yol analizi, değişkenlerin neden-sonuç ilişkisini analiz etmek için kullanılan çoklu doğrusal regresyon analizinin bir türüdür. Karmaşık sistemlerdeki değişkenler arasındaki ilişkileri incelemek için ilk kez 1918 olmak üzere 1920'lerde ve 1930'larda Amerikalı bir genetikçi olan Sewall Wright tarafından geliştirilmiştir[31-33].

Yol analizi, değişkenler arasındaki neden-sonuç ilişkilerini analiz etmek için kullanılan bir çoklu doğrusal regresyon yöntemidir. Bu yöntem, regresyon modelinin bir genişlemesi olarak kabul edilir ve aralarındaki ilişkiler karmaşık

olan değişkenlerin analizinde sıklıkla kullanılır. Yol analizi, değişkenler arasındaki neden-sonuç ilişkilerini bir oklar ve yol katsayıları ile gösteren bir yol diyagramı kullanır. Bu yöntem, yapısal eşitlik modellemesinin gelişmesine önayak olmuştur.

YEM yolu modelleme yaklaşımları, genetikçi Wright'ın fikirleri üzerine inşa edilerek Cowles Komisyonu'nda geliştirilmiş ve Cowles'ta Nobel uzmanı Trygve Haavelmo tarafından desteklenmiştir. Cowles Komisyonu, ekonomide yapısal modellerin geliştirilmesinde önemli bir rol oynamıştır. Komisyon, 1930'larda bir grup ekonomist tarafından kurulmuş ve ekonometri yöntemlerini geliştirmek için birçok önemli çalışma yapmıştır[34]. Yapısal modellerin geliştirilmesinde önemli bir adım, Trygve Haavelmo'nun 1943 tarihli makalesinde atıfta bulunarak yapılan çalışmalarla atılmıştır[35]. Haavelmo, Cowles Komisyonu'nda çalışan bir ekonomisttir ve yapısal modellerin geliştirilmesindeki önemli katkılarından dolayı 1989 Nobel Ekonomi Ödülü'ne layık görülmüştür[36].

Haavelmo'nun 1943 tarihli makalesinde, ekonometrik modellerin birleştirilmesi ve özellikle birleşik yapısal eşitlik modellemesi üzerinde durulmuştur[35]. Bu model, YEM'in temelini oluşturur. Haavelmo, yapısal eşitlik modellemesi kavramını ilk kez kullanmamış olsa da, makalesinde bu kavramı daha yaygın bir şekilde kullanmaya başlamış ve modelin geliştirilmesinde önemli bir rol oynamıştır.

Goldberger'in çalışmaları ve *Structural Equation Models in Social Sciences* başlıklı yayın, YEM modellerine olan ilginin artmasına katkıda bulunmuştur[37-39].

Ancak pratikte, yapısal modellerin gerçek gelişimi, istatistik yazılımlarının dinamik gelişimine ve ölçme modellerinin yapısal modellerle birleştirilmesin ve daha sonra da psikometrik ve ekonometri alanında genişlemesine bağlanabilir[40].

YEM ile ilgili metodolojik kavramlar, ilginç bir şekilde eş zamanlı ve fakat birbirlerinden bağımsız bir şekilde Jöreskog [41, 42], Keesling [43] ve Wiley [39] çalışmalarında olarak önerilmiş olmasına rağmen Jöreskog literatürde ilk YEM modelini ve YEM modeli kurulmasına olanak sağlayan LISREL yazılımını geliştirmesiyle tanınmıştır[42].

LISREL, analiz edilecek kovaryans yapılarının genel bir bilgisayar programı olan ACOVS'un Jöreskog, Gruavaeus ve van Thillo ile birlikte önceden icat ettiği bir programın bir öncüsü niteliğindeydi[42, 44].

Ayrıca, LISREL üzerindeki çalışmalar, Jöreskog'un Princeton Üniversitesi Eğitim Testi Servisi'nden İsveç Uppsala Üniversitesi'nde istatistik profesörlüğüne geçtiği 1972 yılında başladı. Akademik meslektaşı Sörbom, LISREL'deki tüm programlama şemalarını hazırladı ve gizli ortalamaların tahmin edilmesi yöntemini geliştirdi. Ancak, çalışmalarına başlamadan önce Jöreskog, o dönemde Wright ve diğer sosyologların yaptığı yol analizi ve moment tahminciler, psikometriden faktör modelleri ve istatistik ve ekonometriden etkili tahmin ve Neyman-Pearson hipotez testi ile ilgili olarak çapraz disipliner entegrasyonun bir örneği olan Hauser ve Goldberger'ın kitap bölümünden [45] yararlandı. Hauser ve Goldberger, sınırlı bilgi tahmini teorisine odaklandılar ve maksimum olabilirlikle tahmin edilen yapısal denklemler model sistemi arkasındaki gerçekleri açığa çıkarmaya çalıştılar. Jöreskog[42], yaklaşımlarını özetledi, Yapısal Eşitlik Modellerinin tahmin edilmesi için maksimum olabilirlik çerçevesini sunmuş ve yukarıda bahsedilen LISREL bilgisayar yazılımını ampirik uygulamalar için geliştirdi.

LISREL ile birlikte, YEM'in yazılım geliştirme alanındaki gerçek patlama, şu bilgisayar paketleriyle gerçekleşmiştir.

- EQS[46],
- LISCOMP, yeni adı ile MPLUS [47, 48],
- AMOS [49],
- PROC CALIS (in SAS),
- HLM [50],
- SIMPLIS [51]
- GLAMM [52, 53],
- R (açık kaynak) istatistik ortamıyla ilgili ücretsiz yazılım paketleri,
- SEM package [54],
- LAVAAN [55]

Tüm bu yazılımların ortak avantajı, deney tasarımlarının simülasyonunu yaparak ve analiz edilen değişkenler arasındaki korelasyonları daha doğru bir şekilde doğrulayarak, sebep-sonuç ilişkilerinin test edilebilirliğini sağlayan son derece gelişmiş ve hızlı hesaplama çözümleri sunmasıdır.

Yukarıdaki yazıda belirtilen yazılım gelişmeleri döneminde, hesap makineleri ve analog hesaplama yöntemleri gibi önceden mevcut olan yöntemlerin sınırlarını aşan önemli ölçüde otomatik hesaplama yöntemlerinde önemli gelişmeler kaydedilmiştir. Bilgisayar teknolojisindeki bu gelişmeler, hem uzmanların hem de yeni başlayanların genellikle karmaşık ve yapılandırılmamış problemlere dayanan büyük veri setlerine dayanan yoğun analizleri yapmalarına olanak tanımıştır[56]. Bunun bir sonucu olarak, Hershberger, yapısal eşitlik modellemesi ile ilgili bilimsel makale ve yayın sayısının 1994-2001 yılları arasında arttığını belirtmiştir[57].

2.5 Yapısal Eşitlik Modellemesine Ait Yaklaşımlar

Yapısal eşitlik modellemesinde iki genel yaklaşım vardır: LISREL, AMOS ve EQS'de programlarında kullanılan kovaryans temelli yapısal eşitlik modellemesi (K-YEM / CB-SEM) ve kısmi en küçük kareler yapısal eşitlik modellemesi (KEKK-YEM / PLS-SEM) olarak bilinen varyans temelli yapısal eşitlik modellemesi.

K-YEM, bir dizi yapısal denklem sistemince ima edilen teorik kovaryans matrisinin, model içinde gözlenen ampirik kovaryans matrisine mümkün olduğunca yakın olacak şekilde bir dizi model parametresinin tahmin edilmesine odaklanır. Maksimum olabilirlik yöntemi kullanılarak uygun hale getirildiğinde, bu tahmin, gözlenen göstergelerin çok değişkenli normal dağılımı ve yeterli örneklem büyüklüğü gibi bir dizi varsayımın yerine getirilmesini gerektirir. Bu varsayımlar ihlal edilirse, araştırmacılar için KEKK-YEM uygun bir seçenek olacaktır[58, 59]. K-YEM'nin aksine, KEKK-YEM analizi herhangi bir dağılımsal varsayım yerine getirmeyi gerektirmez [59] ve nispeten küçük örneklem büyüklükleri için sağlam ve doğru uyumlar sağlayabilir[60].

2.5.1 Kovaryans Temelli Yapısal Eşitlik Modellemesi

K-YEM son birkaç on yılda sosyal bilimler alanında yaygın olarak uygulanmış ve hâlâ bugün, özellikle örneklem büyüklüğü büyük, veriler normal dağılım

gösteriyor ve en önemlisi model doğru şekilde belirlenmiş durumdaysa, hipotez testi yoluyla teorilerin doğrulanması veya reddedilmesi için tercih edilen bir veri analizi yöntemidir. Yani, uygun değişkenler, bir teorinin yapısal denklem modeline dönüştürülmesi sürecinde birbirine bağlanır[8, 61, 62]. Ancak, birçok endüstri uygulayıcısı ve araştırmacı, gerçekte bu gereksinimleri karşılayan bir veri seti bulmanın sık sık zor olduğunu belirtmektedir. Ayrıca, araştırma amacı keşif niteliğinde olabilir, yani değişkenler arasındaki ilişkiler hakkında çok fazla bilgi sahibi olunmadığı durumlarda PLS yöntemini düşünebilirler[5].

2.5.1.1 K-YEM Yönteminin Avantajları ve Dezavantajları

K-YEM, sosyal ve davranışsal araştırmalar için birkaç avantaj ve dezavantaja sahiptir. E-YEM'in başlıca avantajlarından biri, birden çok gizil değişken ve gözlenen göstergelerle karmaşık teorik modelleri aynı anda modellenebilmesidir. Bu, karmaşık teorik modellerin test edilmesi ve doğrulanması için güçlü bir araç yapar[63]. Ek olarak, E-YEM, teorik model ile gözlenen veriler arasındaki uygunluk derecesini değerlendirmek için kullanılabilen birkaç uyum indeksi sağlar, araştırmacıların modellerini değerlendirmelerini ve iyileştirmelerini sağlar[64].

E-YEM'in bir diğer avantajı, araştırmacıların ölçüleri güvenilir ve geçerli hale getirmek ve amaçlanan yapıları doğru şekilde ölçtüklerinden emin olmak için ölçüm modelleri belirleyebilmeleridir[65]. Ek olarak, E-YEM eksik verilerle başa çıkabilir ve ölçüm hatasını hesaba katabilir, sonuçların doğruluğunu ve güvenirliğini artırabilir[66].

Ancak, K-YEM'in de birkaç dezavantajı vardır. Başlıca kısıtlamalardan biri, küçük örneklem büyüklüklerinde kullanımının zor olmasıdır[63]. Ek olarak, E-YEM, araştırmacıların güçlü bir teorik altyapıya ve istatistiksel teknikler hakkında bilgi sahibi olmalarını gerektirdiğinden, karmaşık ve teknik bir yöntemdir[67].

2.5.1.2 K-YEM Yönteminin Uygulama Alanları

Kline'a göre K-YEM, pazarlama, psikoloji, sosyoloji ve yönetim gibi alanlarda karmaşık teorik modellerin test edilmesi ve doğrulanması için yaygın olarak kullanılmaktadır[63]. Örneğin, K-YEM, pazarlama araştırmalarında tüketici davranışı modellemesi, marka sadakati araştırması ve pazar bölümlendirmesi

analizi gibi alanlarda kullanılmıştır. Psikolojide, K-YEM, kişilik, zeka ve ruh sağlığı ile ilgili teorik modellerin test edilmesinde kullanılmıştır[68]. Ayrıca, K-YEM, sosyolojide sosyal ağları modellemek, sosyal sermayeyi araştırmak ve sosyal faktörlerin bireysel davranışlar üzerindeki etkisini incelemek için kullanılmıştır[69]. Yönetimde K-YEM, örgütsel davranış, liderlik ve stratejik yönetimle ilgili teorik modellerin test edilmesinde kullanılmıştır[70].

2.5.1.3 K-YEM Yöntemi Nasıl Uygulanır?

Genel olarak, K-YEM'nin uygulanması, model belirleme, tahmin etme ve model uyumunun değerlendirilmesi gibi birkaç adımı içerir ve özel yazılım ve istatistik uzmanlığı gerektirir ve AMOS veya Mplus gibi özel yazılım programları kullanılarak gerçekleştirilebilir[63].

K-YEM'nin uygulanmasının ilk adımı, gizil değişkenleri, gözlemlenen göstergeleri ve aralarındaki ilişkileri içeren bir teorik modelin belirlenmesidir[65]. Araştırmacılar daha sonra, gözlemlenen göstergelerin kovaryans matrisini analiz ederek ve maksimum olabilirlik tahminini kullanarak gizil değişkenler arasındaki ilişkileri tahmin ederek modeli tahmin ederler[36].

Model tahmin edildikten sonra, araştırmacılar modelin verilere ne kadar iyi uyduğunu, ki-kare istatistiği, Karşılaştırmalı Uygunluk İndeksi (CFI) ve Kök Ortalama Kare Hatası (RMSEA) gibi çeşitli uyum indeksleri kullanarak değerlendirirler[64].

Eğer model verilere iyi uyum sağlamazsa, araştırmacılar yol ekleyerek veya çıkararak veya ölçme modelinin belirtiliş şeklini değiştirerek modeli yenileyebilirler.

2.5.1.4 K-YEM'in Geleceği

Kovaryans temelli yapısal eşitlik modellemesi, sosyal ve davranışsal araştırmalarda önemli bir rol oynamaktadır ve geleceği umut vadetmektedir. K-YEM'deki son gelişmeler arasında, daha esnek modelleme ve eksik verilerin ele alınmasına izin veren Bayesci tahminin kullanımı[71], örneklem büyüklüğüne ve normal dağılmayan veriye daha az duyarlı olan Ağırlıklı Kök Ortalama Kare Hatası (WRMR) gibi yeni uyum indekslerinin geliştirilmesi yer almaktadır[72]. Bu

nedenle, K-YEM'nin geleceği umut vadetmektedir ve sürekli gelişimi ve geliştirilmesi, sosyal ve davranışsal araştırmalardaki değerini ve uygulanabilirliğini daha da artıracaktır.

2.5.2 Varyans Temelli Yapısal Eşitlik Modellemesi

KEKK, veri dağılımı hakkında herhangi bir varsayımı olmayan bir YEM yumuşak modelleme yaklaşımıdır. Bu nedenle, KEKK-YEM birçok araştırmacı için K-YEM'e alternatif olarak iyi bir seçenek haline gelir. Gerçekte, PLS, özellikle sınırlı sayıda gözleme sahip ve veri dağılımı çarpık olan, çok uluslu CEO'ları araştırmak gibi uygulamalı araştırma projelerinde yapısal eşitlik modellemesi için faydalı bulunmuştur. KEKK-YEM, davranış bilimleri, pazarlama, organizasyon, yönetim bilgi sistemi ve iş stratejisi gibi birçok alanda kullanılmıştır[5]. Aşağıda iki yöntemin karşılaştırılması maddelenmiştir.

• İstatistiksel yöntemleri farklı olduğu için K-YEM ve KEKK-YEM arasında bir üstünlük yoktur.

• K-YEM'de model uyum ölçütleri, KEKK-YEM'de ise yoktur. K-YEM'de ki-kare, RMSEA, GFI, CFI gibi model uyum ölçütleri kullanılır.

• K-YEM yalnızca ortak varyansları kullanırken, KEKK-YEM toplam varyansı kullanan keşfedici faktör analizi benzeri bir yöntem kullanır.

• KEKK-YEM bir veya iki gösterge ile analiz yapabilirken, K-YEM'de en az üç gösterge gereklidir ve daha az sayıda olması tanımlama sorununa neden olur.

• K-YEM, çok değişkenli normal dağılım varsayımını gerektirirken, KEKK-YEM parametrik olmayan bir analiz yöntemi kullanır.

• KEKK-YEM küçük örneklem büyüklükleri ile analiz yapılabilirken, K-YEM'de bu mümkün değildir.

• Karmaşık modellerde KEKK-YEM daha başarılı olabilirken, K-YEM bu durumda sorun yaratabilir.

• K-YEM sadece reflektif değişkenlerle analiz yapabilirken, KEKK-YEM hem reflektif hem de formatif değişkenleri analiz edebilir[46].

2.5.2.1 KEKK-YEM'in Doğuşu

1960'ların ortalarında, tanınmış bir ekonometrist ve istatistikçi olan Herman Wold, "kısmi en küçük kareler" olarak adlandırılan bir tahmini nedensel sistem kavramını geliştirdi. Bu yöntemi bir varyans tabanlı yapısal eşitlik modeli yaklaşımı olarak, ana bileşen ve kanonik korelasyon analizlerini LISREL veya AMOS'dan daha fazla esneklikle keşifsel modelleme için genişletti. PLS, model parametrelerinin hesaplanmasında dağılımsal varsayımlar yapmazken, basit en küçük kareler çoklu regresyon kullanır ve yumuşak modelleme yaklaşımı olarak bilinir. Bu metot, modelin her parçasını ayrı ayrı uydurduğu (fitting) için gereken örneklem sayısını azaltabilir ancak bununla beraber daha büyük bir örneklemin daha iyi parametre tahmini yapmaya yardımcı olacağı unutulmamalıdır. PLS, gizil değişkenler ve gözlemlenebilir değişkenler arasındaki ilişkilerle ilgilenen dış ölçüm modeline odaklanır. KEKK'nın temel amacı, gizil değişkenlerin tahmini için skor değerlerini elde etmektir ve bu, gizil değişkenlerin tam ağırlıklı doğrusal kombinasyonları olarak gözlemlenebilir değişkenlerin hesaplanması yoluyla yapılır. Bu metodolojinin "kısmi" en küçük kareler olarak adlandırılmasının sebebi, prosedürünün yinelemeli prosedüre sahip olması yani parametreleri aynı anda tahmin etmek yerine ayrı ayrı ele almasıdır. Önemli örnekleme prosedürleri bootstrap, jackknife ve blindfolding olarak sayılabilir.

Bootstrapping, orijinal veri kümesinden yeniden örnekleme yaparak ve her örneklem kullanılarak KEKK modelini yeniden hesaplayarak çoklu örneklem oluşturmayı içerir. Bu işlem tekrarlandığında, parametre tahminlerinin dağılımını elde ederek değişkenlik ve güven aralıklarını değerlendirmek mümkündür[73].

Jackknifing, bir başka yeniden örnekleme tekniğidir ve her seferinde bir gözlemi çıkararak KEKK modelini yeniden tahmin etmeyi ve çıkarılan gözlemi tahmin hatasını hesaplamak için kullanmayı içerir. Bu işlem, tüm gözlemler için tekrarlandığında, tahmin hatalarının bir dağılımını elde etmek ve modelin kararlılığını ve dayanıklılığını değerlendirmek mümkündür[74].

Son olarak blindfolding ise verileri eğitim ve test kümelerine rastgele bölen jackknifing'in bir türevidir. KEKK modeli eğitim kümesi kullanılarak tahmin edilir ve test kümesi kullanılarak tahmin hatası hesaplanır. Bu işlem tekrarlandığında,

modelin yeni verilere genelleme yapabilme yeteneğini değerlendirmek ve tahmin hata dağılımını elde etmek mümkündür[75].

Bu yeniden örnekleme teknikleri, KEKK modellerinin kararlılığı, geçerliliği ve genelleştirilebilirliğini değerlendirmek için önemlidir ve kullanımı, R^2 ve Q^2 gibi diğer uyum indeksleri ile birlikte önerilmektedir[60].

PLS, 50 yılı aşkın bir süre önce geliştirilmesine rağmen, KEKK yazılımı ve dokümantasyonunun eksikliği nedeniyle akademik toplumun dikkatini ancak 1990'ların sonunda çekebilmiştir. Ancak PLS-Graph, VisualPLS, SmartPLS, WarpPLS ve ADANCO gibi grafiksel KEKK yazılımlarının piyasaya sürülmesi ile durum önemli ölçüde iyileşmiştir. İlk uluslararası KEKK konferansı 1999'da yapılmış ve Springer tarafından ilk KEKK rehberi 2010'da yayınlanmıştır. KEKK yönteminin en üst düzey, hakemli dergi makalelerinde, pazarlama ve davranış bilimleri gibi alanlarda kullanımı, her geçen yıl daha fazla artış göstermektedir.

2.5.2.2 KEKK-YEM'in Avantajları ve Dezavatanjları

KEKK yol modelleme yaklaşımının faydaları üzerine önemli miktarda araştırma yayınlanmıştır[62, 76, 77]. Bu faydalar aşağıdaki gibi listelenebilir.

- Küçük örneklem boyutu esnekliği
- Daha az olasılıksal hipotezler
- Değişkenlerin dağılımı hakkında herhangi bir varsayımda bulunmama
- Hata terimlerinin normal dağılmama, heteroskedastisite ve otokorelasyonuna karşı duyarsızlık
- Parametre tanımlama sorunu olmaması
- Gözlemlerin bağımsız olma gerekliliği olmaması
- Gizil değişken ile belirgin değişkenleri hem formasyonel hem de reflektif yollarla araştırma imkânı
- Düzenleyici etki analizi yapma ve potansiyel düzenleyicilerin tanımlanması açısından etkililik
- Hem genel hem bireysel gözlemler için skor üretimi

• Büyük model karmaşıklığına (100 gizil ve 1000 gözlenen değişkene kadar) uygunluk

• Uygun olmayan veya yakınsamayan sonuçların elde edilebileceği araştırmalar için uygunluk

KEKK-YEM'nin Zayıf Yönleri ise şunlardır[5]:

• Küçük örneklem boyutları kullanıldığında yüksek değerli yapısal yol katsayılarının gereksinimi

• Çoklu doğrusallık problemiyle iyi başa çıkamama

• Yönsüz korelasyonun modellenmesi için yöntem sağlamama

• Gizil değişkenlerdeki puanların tam tutarlılığındaki eksiklik nedeniyle bileşen yükleri ve yol katsayıları için önyargılı tahminler üretme olasılığı

• Yük tahminleri ve yol katsayısı tahminleri için büyük ortalama kare hatası oluşturma olasılığı.

2.5.2.3 KEKK-YEM'de Gerekli Minimum Örneklem Büyüklüğü

KEKK-YEM yönteminin küçük örneklemlerde dahi istatistiki olarak güçlü sonuçlar verdiği bilinmektedir[8]. Bununla beraber literatürde bir KEKK-YEM modeli için ne kadar örneklem gerektiğine dair genel bir uzlaşı bulunmamaktadır. En çok tercih edilen yöntemlerden biri 10 katı kuralıdır[78]. Bu kurala göre, örneklem büyüklüğü kurulan modelde bulunan gizil değişken sayısının 10 katı kadar olmalıdır. Diğer taraftan Olya daha genel bir çerçeve çizerek 30 ile 100 arasında bir örneklem büyüklüğünün KEKK-YEM modelleri için yeterli olduğunu belirtmektedir[79]. Bu çalışmanın uygulama kısmında da örneklem büyüklüğü için eşik değeri 30 olarak kabul edilmiştir.

2.6 Gizil ve Gözlemlenen Değişkenler

Gizil değişkenler doğrudan gözlenemeyen, ancak dolaylı olarak ölçülebilen kavramları temsil eder. Bu değişkenler araştırmacının ilgi duyduğu soyut kavramları yakalar. Örneğin, motivasyon, memnuniyet, algılanan kalite gibi gizil değişkenler, gözlemlenen değişkenlerin altında yatan faktörleri temsil eder. Gözlemlenen değişkenler ise araştırmada doğrudan gözlenebilen ve ölçülebilen

değişkenlerdir. Bu değişkenler aracılığıyla elde edilen veriler somut ve açıktır. Örneğin, yaş, cinsiyet, gelir gibi gözlemlenen değişkenler doğrudan ölçülebilir ve kolayca anlaşılabilir.

Gözlemlenen değişkenler, araştırmalarda doğrudan kullanılabilirken, gizil değişkenlerin ölçümü ve analizi daha karmaşık olabilir. Gizil değişkenler genellikle çok sayıda gözlemlenen değişkenin ortak bir faktör altında birleştiği yapısal eşitlik modellemesi gibi analiz teknikleriyle incelenir.

Gözlemlenen ve gizil değişkenler arasındaki ilişkiyi anlamak, araştırmacılara derinlemesine bir anlayış sağlar ve teorik modellerin test edilmesine ve açıklanmasına yardımcı olur. Bu şekilde, araştırmacılar araştırmalarının konusuna daha kapsamlı bir şekilde yaklaşabilir ve sonuçları daha derinlemesine değerlendirebilir.

2.6.1 Gizil Değişken

Gizil (gizli, latent) değişkenler, bir araştırma modelinde yer alan, doğrudan gözlemlenemeyen ve ölçülemeyen kavramları temsil eden değişkenlerdir. Gizil değişkenler, araştırmacının ilgi duyduğu teorik yapıları ve altında yatan ilişkileri açıklamaya çalışır. Bu kavramlar genellikle araştırma alanına özgü olup, örneğin kişilik özellikleri, öğrenme motivasyonu veya marka sadakati gibi kavramlar gizil değişkenlere örnek olarak gösterilebilir[80].

Gizil değişkenlerin özellikleri aşağıda şu şekilde maddelenmiştir.

- Gözlemlenememe: Gizil değişkenler doğrudan gözlenemeyen kavramları temsil ettiği için, araştırmacılar tarafından dolaylı olarak ölçülmeleri gerekmektedir.
- Çoklu gözlemlenenlerin bileşimi: Gizil değişkenler, birden çok gözlemlenen değişkenin bir araya gelmesiyle oluşur ve bu gözlemlenen değişkenlerin altında yatan ortak varyansı yakalar.
- Teorik yapıyı temsil etme: Gizil değişkenler, araştırmacının ilgi duyduğu teorik yapıları ifade eder ve aralarındaki ilişkileri açıklamaya yardımcı olur.

- Ölçeklerin güvenilirliği: Gizil değişkenler, birden çok gözlemlenen değişkenin bir araya gelmesiyle oluştuğu için, bu ölçekler arasındaki rastgele hatalar birleştirilerek ölçüm hatasını azaltır.
- Ölçümlerin eksikliği: Gizil değişkenler, ölçümlerin eksik olduğu durumlarda da kullanılabilir ve eksik veri problemini yönetme esnekliği sağlar[81].

2.6.2 Gözlemlenen Değişken

Gözlemlenen (gösterge) değişkenler, doğrudan gözlenebilen ve ölçülebilen değişkenlerdir. Bunlar, araştırmacının elde ettiği somut verilere dayanır ve genellikle anketler, ölçekler veya gözlem kayıtları yoluyla elde edilir. Örneğin, yaş, cinsiyet, gelir düzeyi gibi demografik değişkenler veya ölçek puanları, öğrencilerin notları gibi veriler, gözlemlenen değişkenlere örnek olarak verilebilir[78, 82].

Gözlemlenen değişkenlerin özellikleri aşağıda sıralanmıştır.

Gözlemlenen değişkenler, araştırmacılar tarafından doğrudan gözlemlenerek veya ölçülerek elde edilen veriler üzerinden analiz edilir[83]. Bu değişkenler, araştırmacının ilgi alanına bağlı olarak farklı ölçüm ölçeklerine sahip olabilir. Örneğin, nominal, ordinal, aralık veya oran ölçeklerinde ölçülebilirler.

Gözlemlenen değişkenler, araştırmada kullanılan veri toplama araçları ve yöntemlerine bağlı olarak değişebilir. Örneğin, anketler, mülakatlar, gözlemleme gibi yöntemlerle toplanan verilerden gözlemlenen değişkenlere ulaşılabilir[83].

Gözlemlenen değişkenler, araştırmacılar tarafından kolaylıkla anlaşılabilir ve yorumlanabilir. Araştırmada kullanılan istatistiksel analizler genellikle gözlemlenen değişkenlere uygulanır ve bu değişkenlerin dağılımı, ilişkileri ve etkileri incelenir[84].

2.6.3 KEKK-YEM'de Gizil Değişkenlerin Gözlemlenen Değişkenlerle İlişkisi

Kısmi En Küçük Kareler Yapısal Eşitlik Modellemesinde, gizil değişkenlerin gözlemlenen değişkenlerle ilişkisi modelin temelini oluşturur. Gizil değişkenler, gözlemlenen değişkenlerin altında yatan kavramsal yapıları temsil eder ve onların varyansını açıklamak için kullanılır[85].

Gizil değişkenler, gözlemlenen değişkenler arasındaki ilişkileri açıklamak ve modelin uyumunu değerlendirmek için kullanılır. Bu ilişkiler, gizil faktörlerin etkileşimini ve neden-sonuç ilişkilerini ifade edebilir[86].

KEKK-YEM'de, gizil değişkenler ve gözlemlenen değişkenler arasındaki ilişkiler genellikle faktör yükleri veya yol katsayıları olarak ifade edilir. Faktör yükleri, gizil değişkenin gözlemlenen değişkeni ne kadar iyi açıkladığını gösteren bir ölçüdür. Yol katsayıları ise bir değişkenin diğer bir değişkeni ne kadar etkilediğini gösterir[85].

KEKK-YEM'de, gizil değişkenlerin gözlemlenen değişkenlerle ilişkisi genellikle korelasyon analizi veya yapısal modelleme yöntemleriyle değerlendirilir[86]. Bu analizler, gizli değişkenlerin gözlemlenen değişkenlerle olan ilişkilerini nicel olarak ölçer ve istatistiksel olarak değerlendirir.

2.7 KEKK-YEM'de Ölçüm Modeli

KEKK-YEM'de ölçüm modeli (dış model), gizil değişkenlerin gözlemlenen değişkenler aracılığıyla ölçülmesini sağlar. Bu aşamada, ölçüm modellemesi, gözlemlenen değişkenlerin gizil değişkenlerle ilişkisini ve gözlemlenen değişkenlerin birbirleriyle ilişkisini belirlemek için kullanılır.

Ölçüm modeli genellikle faktör analizi veya bileşen analizi gibi yöntemlerle gerçekleştirilir. Bu analizler, gözlemlenen değişkenlerin birbirleriyle olan ilişkisini açıklamak için faktörler veya bileşenler aracılığıyla yapısal bir model oluşturur. Bu şekilde, gözlemlenen değişkenlerin gizil değişkenlerin altında yatan kavramsal yapıyı yansıttığı kabul edilir.

KEKK-YEM'de ölçüm modellemesi için farklı yaklaşımlar mevcuttur. Faktör analizi temelli yaklaşımlar, gözlemlenen değişkenler arasındaki kovaryans veya korelasyon yapılarını analiz ederken, bileşen analizi temelli yaklaşımlar, gözlemlenen değişkenlerin varyansını açıklamak üzerine odaklanır.

Ölçüm modellemesi aşamasında, ölçek uyumunu değerlendirmek için çeşitli istatistiksel ölçütler kullanılır. Bu ölçütler arasında faktör yükleri, kapsamalar, korelasyonlar ve iç tutarlılık gibi ölçümler yer alır.

2.7.1 Yapısal Eşitlik Modelinde Ölçüm Modelinin Önemi

Yapısal eşitlik modelinde ölçüm modeli, araştırmada kullanılan gözlemlenen değişkenlerin, altında yatan gizil veya kavramsal yapıları doğru bir şekilde yansıtması için büyük bir öneme sahiptir[78]. Ölçüm modeli, gözlemlenen değişkenlerin gizil değişkenlerle ilişkisini ortaya koyarak, bu ilişkilerin ölçülmesini sağlar ve yapısal modelin güvenilir ve geçerli bir şekilde değerlendirilmesini sağlar.

Ölçüm modeli, yapısal eşitlik modelinin temelini oluşturur ve modelin doğru bir şekilde ölçüldüğünden emin olmayı sağlar[78]. Yanlış ölçüm modeli, model sonuçlarında ciddi sapmalar ve hatalara yol açabilir. Bu nedenle, ölçüm modelinin doğru ve güvenilir bir şekilde oluşturulması, yapısal eşitlik modelinin geçerliliği ve güvenilirliği için kritik bir adımdır.

Ölçüm modeli, faktör analizi veya bileşen analizi gibi istatistiksel yöntemlerle gerçekleştirilebilir[78, 87]. Bu yöntemler, gözlemlenen değişkenlerin birbirleriyle olan ilişkisini ve altında yatan faktör veya bileşen yapılarını açıklamak için kullanılır. Ölçüm modelinin doğru bir şekilde oluşturulması, araştırmacılara gizil değişkenlerin ölçümlerini yapma ve analiz etme imkânı sağlar.

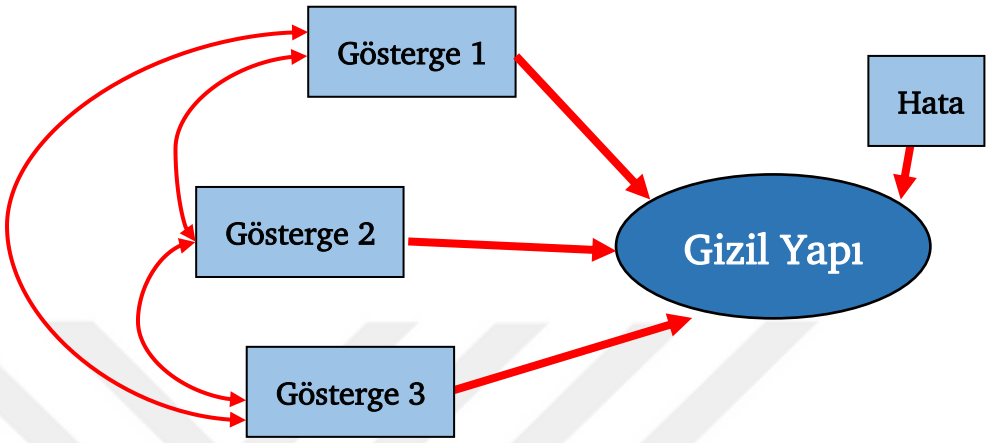
Ölçüm modeli aşamasında, çeşitli uyum ölçütleri kullanılarak modelin uyumunun değerlendirilmesi önemlidir[86]. Örneğin, faktör yükleri, kapsamalar, korelasyonlar ve iç tutarlılık gibi istatistiksel ölçümler, ölçüm modelinin doğruluğunu ve güvenilirliğini değerlendirmek için kullanılır.

2.7.2 Formatif Ve Reflektif Ölçeklerin Kullanımı

Formatif (biçimlendirici/nedensel) ve reflektif (yansıtıcı) ölçekler, kısmi en küçük kareler yapısal eşitlik modellerinde kullanılan iki farklı ölçüm yaklaşımıdır[87]. Bu ölçekler, gözlemlenen değişkenlerin gizil veya kavramsal yapılarıyla ilişkisini açıklamak için kullanılır.

Formatif ölçekler, gözlemlenen değişkenlerin altında yatan gizil değişkeni belirlemek veya oluşturmak için kullanılır[88]. Bu ölçekler, değişkenlerin özelliklerini bir araya getirerek veya belirli bir özelliği temsil ederek gizil değişkeni açıklar. Formatif ölçekler, tek bir yönde kavramsal bir ilişkiyi yansıtabilir ve

değişkenler arasında bir etkileşim olabilir. Ölçek yapısının doğruluğu ve güvenilirliği formatif ölçeklerde daha az önemlidir, çünkü formatif ölçeklerin güvenilirliği önceden belirlenemez.



Şekil 2.3 Formatif yapısal eşitlik modeli

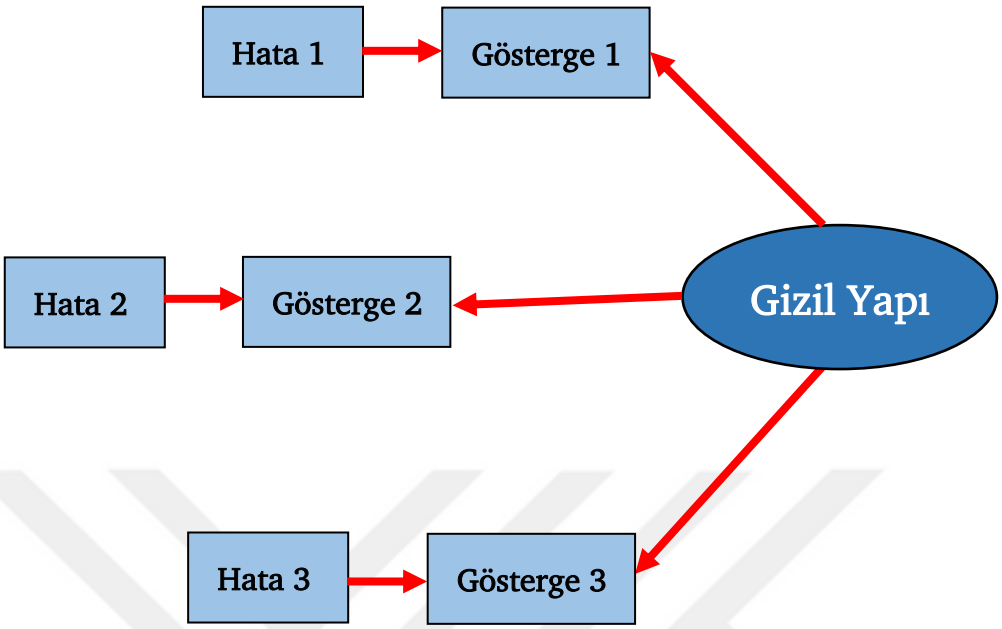
Şekil 2.3'te gösterge adıyla isimlendirilen değişkenler gözlemlenebilen değişkenler iken yapı adıyla ifade edilen değişken ise araştırılan gizil değişkendir. Bu modelde, gözlemlenebilen değişkenlerdeki bir veya birkaç değişiklik gizil değişken de değişikliklere sebep olur. Formatif modellerde göstergeler nedensel göstergelerdir ve kendi aralarında pozitif, negatif ya da sıfır korelasyon gösterebilir[89].

χ_{pm} her bir açıklayıcı değişkeni, ω_{pm} ilişki katsayılarını, ξ_{pm} gizil değişkeni, δ_{pm} gizil değişkenlere ait hatayı göstermek üzere, formatif modeller şu şekilde kurulur.

$$\boldsymbol{\xi_{pm}} = \sum_p \boldsymbol{\omega_{pm}} \cdot \boldsymbol{\chi_{pm}} + \boldsymbol{\delta_{pm}} \tag{2.29}$$

Reflektif ölçekler ise gözlemlenen değişkenlerin gizil değişkenleri yansıttığı düşünceye dayanır[36]. Bu ölçekler, gizil değişkenin altında yatan özellikleri yansıtır ve bu özellikleri ölçer. Reflektif ölçekler, değişkenler arasında bir korelasyonu yansıtabilir ve çoğu zaman çok yönlü bir ilişkiyi ifade eder. Reflektif ölçeklerin güvenilirlik ve geçerlilik testlerine tabi tutulması önemlidir, çünkü bu

ölçeklerin doğru ve güvenilir bir şekilde ölçtüğünden emin olunması gerekmektedir.



Şekil 2.4 Reflektif yapısal eşitlik modeli

χ_{pm} her bir açıklayıcı değişkeni (gösterge), λ_{p0} kesen, λ_{pm} yük katsayılarını, ξ_{pm} gizil değişkeni, ε_{pm} hata terimlerini göstermek üzere, reflektif modeller şu şekilde kurulur.

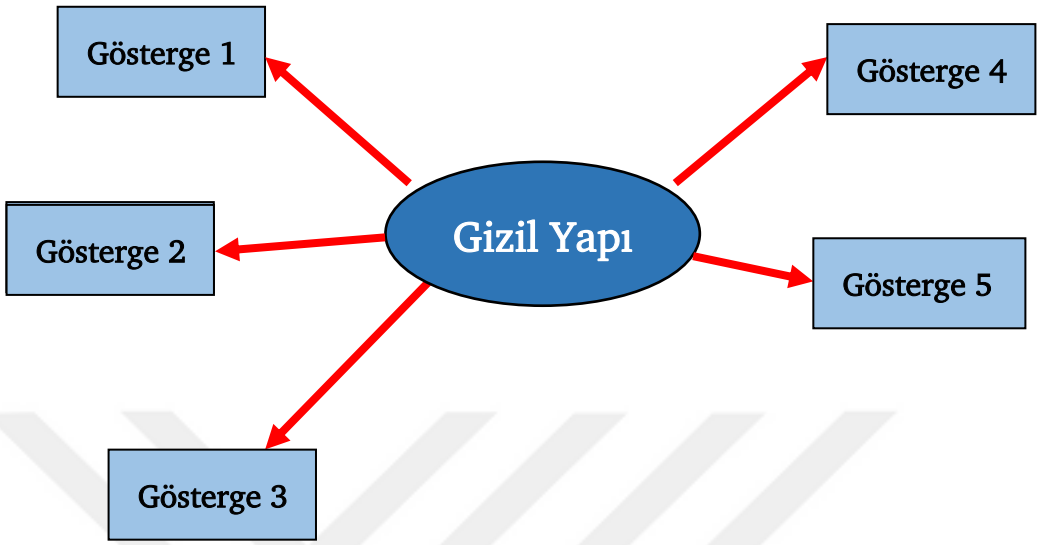
$$\boldsymbol{\chi_{hm} = \lambda_{h0} + \lambda_{hm} \cdot \xi_m + \varepsilon_{hm}} \qquad (2.30)$$

Reflektif modellerde, gizil değişkendeki değişiklikler direkt olarak gözlemlenebilen değişkenlerde de değişimlere sebep olur. Göstergeler yansıtıcıdırlar ve kendi aralarında yüksek korelasyon göstermelidir[89].

KEKK-YEM modellerinde formatif ve reflektif ölçeklerin kullanımı, araştırmanın amacına ve araştırılan konuya bağlıdır[78]. Formatif ölçekler, yeni bir konuyu keşfetmek veya bir teoriyi geliştirmek için kullanılabilirken, reflektif ölçekler genellikle önceden belirlenmiş bir teoriyi test etmek veya mevcut bir yapıyı doğrulamak için tercih edilir.

Çoklu gösterge ve çoklu neden modeli (MIMIC - multiple indicators and multiple causes model) olarak da bilinen model, yansıtıcı ve yapılandırıcı ölçüm

modellerinin bir kombinasyonu olarak düşünülebilir (Vinzi vd., 2010; Banet, 2015).



Şekil 2.5 MIMIC modeli

2.7.3 Gözlemlenen Değişkenlerin Gizil Değişkenlerle İlişkilendirilmesi

Yapısal eşitlik modellemesinde, gözlemlenen değişkenlerin gizil değişkenlerle ilişkilendirilmesi, aralarındaki ilişkileri yakalamak ve ilgi duyulan kavramları ölçmek için kullanılır[78]. Gözlemlenen değişkenler ile gizil değişkenler arasındaki ilişkiler, yol katsayıları ile temsil edilir ve ilişkinin gücünü ve yönünü gösterir.

Gözlemlenen değişkenleri gizil değişkenlerle ilişkilendirmek süreci, faktör yüklerini veya yol katsayılarını tahmin etmeyi içerir ve gözlemlenen değişkenlerin gizil değişkenlerin ölçülmesine katkı düzeyini gösterir[90]. Bu yüklemeler, gözlemlenen değişkenler ile ölçmeyi amaçladıkları temel kavramlar arasındaki ilişkiyi temsil eder.

Gözlemlenen değişkenler ile gizil değişkenler arasındaki ilişkiler, YEM'de oldukça önemlidir çünkü doğrudan gözlemlenemeyen değişkenler arasındaki ilişkileri araştırmacılara inceleme olanağı sağlar[63]. Gözlemlenen değişkenleri gizil

değişkenlerle ilişkilendirmek, araştırmacıların temel teorik yapıları incelemelerine ve ilişkilerine dair hipotezleri test etmelerine olanak tanır.

Gözlemlenen değişkenler ile gizil değişkenler arasındaki ilişkilerin belirlenmesi, araştırmanın teorik temelleri ve kavramsal çerçevesi temelinde yapılması gerektiği unutulmamalıdır[78]. YEM modelinde ilişkilerin belirlenmesi ve tahmin edilmesi aşamasında, araştırmacılar teorik temellere ve deneysel kanıtlara dikkatlice göz atmaları gerekmektedir.

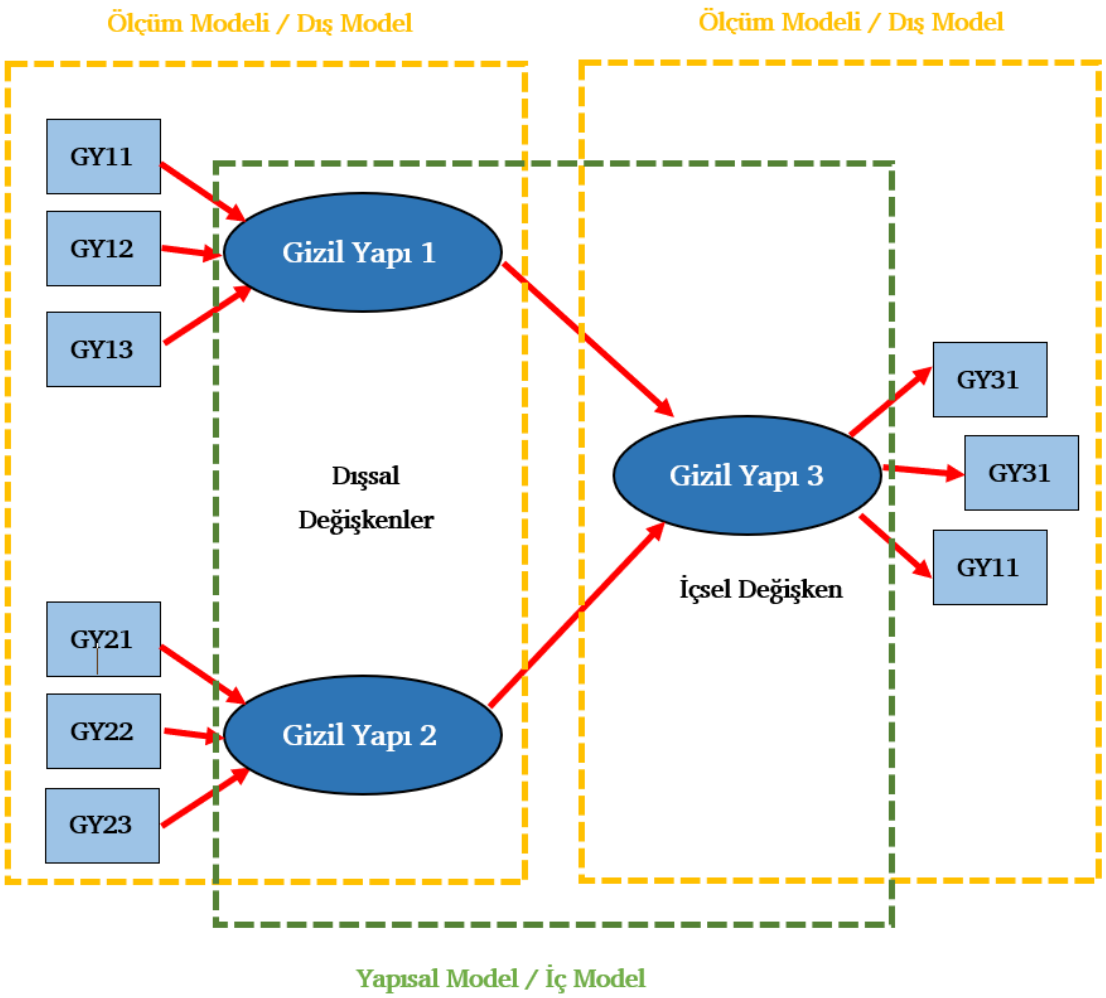
2.8 KEKK-YEM'de Yapısal Model

KEKK-YEM'de yapısal model (içsel model), araştırma çerçevesindeki gizil değişkenler (kurgusal yapılar) arasındaki ilişkileri temsil eden genel modelin bir bileşenidir[78]. Gözlenen belirteçleri doğrudan modellemek yerine, gizil değişkenler arasındaki yapısal ilişkileri ve etkileşimleri yakalamaya odaklanır.

İçsel model, çalışmanın teorik çerçevesine ve araştırma hipotezlerine dayanarak oluşturulur. Bu model, gizil değişkenler arasındaki yönlü ilişkileri temsil eden oklar veya yolaklar içerir ve bunların arasındaki hipotez edilen nedensel veya korelasyonel bağlantıları gösterir. İçsel model genellikle yol diyagramı olarak belirtilir, burada her ok, karşılık gelen gizil değişkenler arasındaki ilişkinin gücünü ve yönünü ölçen yol katsayısını temsil eder.

İçsel model, araştırmacılara gizil değişkenler arasındaki doğrudan ve dolaylı etkileri inceleme ve teorik hipotezlerini test etme imkanı sağlar[78]. Araştırmacılara altta yatan yapısal ilişkileri anlama konusunda yardımcı olur ve gizil değişkenlerin araştırma bağlamında nasıl birbirleriyle ilişkili olduğunu anlamalarına yardımcı olur.

Ayrıca, içsel model, KEKK-YEM'de yol katsayılarını tahmin etmek ve ilişkilerin anlamlılığını değerlendirmek için temel olarak kullanılır. İçsel model, gizil değişkenlerle gözlenen belirteçleri arasındaki ilişkileri temsil eden dış modelle birlikte kullanılarak kapsamlı sonuçlar ve yorumlar elde etmek için kullanılır.



Şekil 2.6 Örnek yapısal model

2.8.1 Egzojen ve Endojen Değişkenler

Yapısal modellerde, egzojen (dışsal) ve endojen (içsel) değişkenler temel kavramlardır. Egzojen değişkenler dışarıdan, başka bir sistemden veya dış etkilerden kaynaklanan ve modelde bağımlı değişkenler üzerinde etkisi olan değişkenlerdir. Endojen değişkenler ise modelde incelenen ve diğer değişkenlerle birlikte ilişkili olan değişkenlerdir[63].

Egzojen değişkenler, modeldeki diğer değişkenlerden etkilenmez ve dışarıdan gelen etkileri yansıtır. Bu değişkenler modelin girdi noktalarını temsil eder ve genellikle araştırmacının kontrolü dışındadır.

Endojen değişkenler ise modelde incelenen ve diğer değişkenlerle ilişkili olan değişkenlerdir. Bu değişkenler, modelde diğer değişkenler tarafından etkilenebilir.

Egzojen ve endojen değişkenler arasındaki fark, etkileme ilişkileridir. Egzogen değişkenler, diğer değişkenler üzerinde etki yapar ancak kendisi diğer değişkenler tarafından etkilenmez. Endojen değişkenler ise diğer değişkenlerden etkilenebilir ve etkileyebilir. Bu nedenle, endojen değişkenler arasında karşılıklı ilişkiler bulunabilir.

Şekil 2.6'da Gizil Yapı 1 ve Gizil Yapı 2, Gizil Yapı 3'ü etkiledikleri için dışsal değişken durumundadır. Gizil Yapı 3 ise diğer değişkenler tarafından etkilenen içsel değişken pozisyonundadır.

Egzojen değişkenler bağımsız değişkenlere, endojen değişkenler ise bağımlı değişkenlere benze de bu kavramlar tam olarak birbirlerine karşılamaz. Çünkü bir değişken yapısal modelde başka değişkenler tarafından tahmin edilirken aynı anda başka değişkenleri tahminliyor (etki ediyor) olabilir.

Yapısal modellerde egzojen ve endojen değişkenlerin belirlenmesi, araştırmacının teorik çerçevesine ve araştırma hipotezlerine dayanır. Egzogen ve endojen değişkenlerin doğru şekilde tanımlanması, modelin yapısal ilişkilerini anlamak ve test etmek için önemlidir.

ξ gizli değişken, β yol katsayıları matrisi ve ξ içsel hata terimi göstermek üzere yapısal eşitlik modeli şöyle gösterilir.

$$\boldsymbol{\xi_m = \beta \xi_m + \zeta_m} \tag{2.31}$$

2.9 KEKK-YEM Algoritması

Jan-Bernd Lohmöller, 1989 yılında KEKK-YEM'e ilişkin bir algoritma geliştirmiştir[91]. KEKK-YEM, dışsal ağırlıkların w_q ve gizli değişken skorunun $\xi_{q(skor)}$ tahmin edilmesini sağlayan bir yinelemeli yöntemdir. Algoritma, blokların tekli ve çoklu regresyon modelleri ile tek tek çözülmesiyle gerçekleşir, bu nedenle tahmin süreci kısmen gerçekleşir. B_j değerleri, yapısal ilişkilere dayalı olarak tahmin edilen gizli değişken skorlarına göre yapılan regresyon analizine dayanarak hesaplanır[92].

P değişken sayısını, Q blok sayısını, N birim sayısını göstermek üzere KEKK-YEM algoritmasının aşamaları aşağıdaki gibidir[92-95].

1. Aşama: Dışsal yaklaşım veya dışsal tahmin, v_q olarak ifade edilir ve her bir gizli değişken, kendi göstergelerinin lineer bir kombinasyonu olarak kestirilir. Göstergeler x_{pq} ile ağırlıkların w_{pq} doğrusal kombinasyonu olarak;

$$\boldsymbol{v_q \propto \pm \sum_{p=1}^{P_q} w_{pq} x_{pq} = \pm X_q w_q} \tag{2.32}$$

eşitliği ile hesaplanır.

2. Aşama: İçsel yaklaşımda veya içsel tahminde, gizil değişkenler komşu (yan yana) giziz değişkenlerin (Q') doğrusal bir kombinasyonu olarak tanımlanır.

$$\boldsymbol{\vartheta_q \propto \sum_{q'=1}^{Q'} e_{qq'} v_{q'}} \tag{2.33}$$

$\boldsymbol{\vartheta_q}$, ξ_q'nın standartlaştırılmış içsel tahmin değeridir. Ayrıca her bir içsel ağırlık $\boldsymbol{e_{qq'}}$, q'. gizil değişkenin dışsal tahmini v_q ile q'. gizil değişkenin dışsal tahmini $\boldsymbol{v_{q'}}$ arasındaki korelasyonun işaretine eşittir.

İçsel ağırlık 3 şekilde hesaplanır:

Merkezi ağırlık şeması: q . gizil değişkenin dışsal tahmini v_q ile q. gizil değişkenin dışsal tahmini $\boldsymbol{v_{q'}}$ arasındaki korelasyonun işareti alınır[96].

$$\boldsymbol{e_{qq'} = \mathbf{sign}\{\mathbf{corr}(v_q, v_{q'})\}} \tag{2.34}$$

Faktöriyel ağırlık şeması: q′. gizil değişkenin dışsal tahmini v_q ile q′. gizil değişkenin dışsal tahmini $v_{q'}$ arasındaki korelasyon alınır[91].

$$\boldsymbol{e_{qq'} = \{\mathbf{corr}(v_q, v_{q'})\}} \tag{2.35}$$

Yapısal şema veya yol ağırlık şeması: Eğer yapısal eşitlikte vq endojen değişken olarak görev yapıyorsa q . gizil değişkenin dışsal tahmini vq ile q′. gizil değişkenin dışsal tahmini vq ′ arasındaki regresyon katsayısı alınır.

Yol ağırlık şeması, gizli değişkenler arasındaki yapısal ilişkilerin yönüyle uyumlu görünse de, göstergeler arasında güçlü bir ilişki olduğunda daha sık kullanılan merkezi ağırlık şeması tercih edilir. Faktöriyel ağırlık şeması ise ilişkilerin zayıf olduğu durumlarda tercih edilir. Farklı şemaların mevcut olmasına rağmen, yapısal modele ilişkin ilişki yönünü dikkate alan tek şema olan yol ağırlık şemasının kullanılması önerilir[93].

3. Aşama: Yeni dışsal ağırlıklar içsel ve dışsal tahminlere göre güncellenir. Yeni dışsal ağırlıklar elde etmek için 3 farklı mod bulunmaktadır.

Mod A: Yansıtıcı olarak, her bir dışsal ağırlık w_{pq} , içsel tahmin $\boldsymbol{\vartheta_q}$ ile gösterge x_{pq} kullanılarak yapılan regresyon analizi sonucunda $\boldsymbol{\vartheta_q}$'nın regresyon katsayısı olarak hesaplanır. $\boldsymbol{\vartheta_q}$ standartlaştırıldığı için dışsal ağırlık w_{pq} ,

$$\boldsymbol{w_{pq} = \mathbf{cov}\left(x_{pq}, \vartheta_q\right)} \tag{2.36}$$

eşitliği ile hesaplanır.

Bu durumda, regresyon katsayısı, her bir göstergenin kendi gizil değişkeninin içsel tahminiyle olan ilişkisini gösteren kovaryansı ifade eder. Göstergeler standartlaştırıldığında, bu kovaryans bir korelasyona dönüşür.

Mod B: Biçimlendirici olarak, içsel tahmin $\boldsymbol{\vartheta_q}$, izil yapı ile ilişkili göstergelerin X_q regresyon analizinin yapılmasıyla hesaplanır. Ağırlıklar w_q ise, $\boldsymbol{v_q = \sum_p w_{pq} x_{pq}}$ olmak üzere;

$$\boldsymbol{w_q = \left(X_q' X_q\right) X_q' \vartheta_q} \tag{2.37}$$

eşitliği ile güncellenir.

Mod C: Lohmöller'in bir versiyonu olarak uygulanır ve Mod B'nin özel bir durumudur. MIMIC yansıtıcı ve biçimlendirici yolların her ikisini de kapsar, bu nedenle h sayıda gösterge için yol katsayısı yansıtıcı olarak basit doğrusal regresyon ile tahmin edilir. h=1 ile p_1 için $\chi_{hq} = p_{hq}\xi_q$ olur ve h=p_1+1 ile P için yol katsayıları hesaplanır. Biçimlendirici yolla bağlanan göstergeler çoklu doğrusal regresyon ile tahmin edilir[92].

$$\xi_q = \sum_{h=p_1+1}^{P_q} w_{hq} x_{hq} \quad (2.38)$$

4. Aşama: Yakınsaklık, dışsal tahmin ve içsel tahmin karşılaştırılır ve yakınsamak zorundadırlar. Eğer yakınsamazlarsa, algortima 1.adımdan itibaren tekrarlanır. Her bir yineleme adımının yakınsaklığı, Z adımın dışsal ağırlıkları ile Z-1 adımın dışsal ağırlıkları karşılaştırılarak kontrol edilir[92]. Yakınsaklık ölçütü, Wold tarafından [97] şöyle önerilmiştir: $|\hat{w}^{Z-1} - \hat{w}^{Z}| < 10^{-5}$

5. Aşama: Katsayıların tahmini, yol katsayıları $B_{qq'}$ ve yük katsayıları $\hat{\lambda}_{qq'}$ içsel ve dışsal modellere göre tahmin edilir.

Yapısal model için yol katsayıları EKK ile tahmin edilir:

$$\xi_q = \sum_{q'} \beta_{qq'} \xi_{q(skor)} \quad (2.39)$$

$$\beta_{qq'} = \left(\xi'_q \xi_q\right)^{-1} \xi'_q \xi_{q(skor)} \quad (2.40)$$

Yansıtıcı ölçüm modeli için yük katsayıları basit doğrusal regresyonun regresyon katsayılarıdır.

$$x_{pq} = \lambda_{pq} \xi_q \quad (2.41)$$

$$\lambda_{pq} = \left(\xi'_q \xi_q\right)^{-1} \xi'_q x_{pq} \quad (2.42)$$

2.10 KEKK-YEM Modellerinin Değerlendirilmesi

KEKK-YEM modellerinin değerlendirilmesi aşağıdaki adımlarla gerçekleştirilir:



Şekil 2.7 KEKK-YEM modellemesinin değerlendirme aşamaları

2.10.1 KEKK-YEM Modellerinde Güvenirlik Ve Geçerlik Kontrolü

KEKK-YEM modellerinin geçerliği ve güvenirliği 3 başlık altında değerlendirilir. Bunlar

- İç Tutarlık Güvenirliği
- Birleşme Geçerliği
- Ayrışma Geçerliği

2.10.1.1 İç Tutarlık Güvenirliği

İç tutarlık güvenilirliği, genellikle Cronbach's Alpha (CA) değeri ile değerlendirilir. Bu ölçü, ölçeğin tüm öğeleri arasındaki korelasyonları ve öğelerin toplam puanı arasındaki korelasyonu dikkate alır[98]. Cronbach'ın Alfa değeri, 0 ile 1 arasında bir değer alır ve yüksek Cronbach'ın Alfa değerleri, ölçeğin iç tutarlılık güvenilirliğinin yüksek olduğunu gösterir. CA'nın Genellikle 0.7 veya daha yüksek olması beklenir[99-101].

K, modeldeki öğelerin sayısını, r ise anlamlı ortalama gösterge korelasyon katsayısı (yani alt veya üst üçgen korelasyon matrisinin ortalaması) olmak üzere;

$$\textbf{Cronbach } \alpha = \frac{\mathbf{K} \,.\, \bar{\mathbf{r}}}{[\mathbf{1} + (\mathbf{K} - \mathbf{1}) \,.\, \bar{\mathbf{r}}]} \tag{2.43}$$

Diğer tarafta KEKK-YEM modeli için, Henseler vd. [102] Cronbach's Alpha yerine rho_A değerini, Hair vd. [100] ise Bileşik Güvenirlik (Composite Reliability - CR) değerini önermektedir. Bu önerinin temel nedeni, Cronbach's Alpha değerinin gerçek güvenilirliği altında hesaplandığı iddiasıdır[103]. KEKK-YEM modellerinde hesaplanan rho_A değerleri için kabul edilir en düşük değer 0.70'ir[102]. CR değerleri için ise kabul edilebilir en küçük seviye bazı kaynaklarda 0.60 [104], bazılarında ise 0.70 [100] olarak belirtilmiştir. Bileşik güvenirlik ρ_c ile gösterilir ve şu şekilde hesaplanır.

$$\boldsymbol{\rho_c} = \frac{\left(\sum_{k=1}^{K} \boldsymbol{l_k}\right)^2}{(\sum_{k=1}^{K} \boldsymbol{l_k})^2 + (\sum_{k=1}^{K} \boldsymbol{var(e_k)})} \tag{2.44}$$

Bu formülde l_k, K göstergeleri ile ölçülen belirli bir yapının gösterge değişkeni k'nin standartlaştırılmış dış yükünü, e_k gösterge değişkeni k'nin ölçüm hatasını, var(e_k) ise ($1-l_k^2$) olarak tanımlanan ölçüm hatasının varyansını gösterir.

Güvenirlik açısından dikkate alınması gereken bir diğer ölçüt ise gösterge güvenilirliğidir. Gösterge güvenilirliği, her bir gösterge yükünün karesi olarak hesaplanır. Gösterge güvenilirliği, bir değişkendeki varyansın göstergeler tarafından açıklanan oranını ifade eder. Bu değer için değer aralığı Hulland'a göre 0.40 ve üzeri [89], Hair ve arkadaşlarına göre 0.50 üzeri [100], Kock ve Lynn'e göre ise 0.70 ve üzeridir.

2.10.1.2 Birleşme Geçerliği

KEKK-YEM modellerinde birleşme geçerliği, yapısal modelin ölçüm modeli ile uyumlu olduğunu ve kullanılan ölçeklerin teorik olarak beklenen faktörleri ölçtüğünü gösteren bir değerlendirme kriteridir. Bu geçerlilik, göstergelerin faktörlerle yükleri arasındaki ilişkilerin istatistiksel olarak anlamlı ve beklenen yönde olduğunu doğrulamayı amaçlar.

Birleşme geçerliği, genellikle göstergelerin yapısal modeldeki faktörlere olan yüklerinin değerlendirilmesiyle değerlendirilir. Bu değerlendirme, göstergelerin faktörlerle güçlü bir ilişkiye sahip olup olmadığını belirlemek için faktör yüklerinin

değerlendirilmesini içerir. Yüksek ve istatistiksel olarak anlamlı faktör yükleri, göstergelerin faktörleri iyi ölçtüğünü gösterirken, düşük veya istatistiksel olarak anlamsız yükler birleşme geçerliğini olumsuz etkileyebilir.

Araştırmacılar, birleşme geçerliğini değerlendirmek için genellikle faktör yüklerini kontrol etmektedirler. Faktör yüklerinin değerlendirilmesinde, genellikle kabul edilebilir eşik değeri belirlenir 0.60'tır[105].

Birleşme geçerliliği için kabul gören bir parametre açıklanan ortalama varyans (AVE) değeridir. AVE değeri, bir faktörün göstergelerle güçlü bir şekilde ilişkili olduğunu ve göstergelerin faktörü iyi ölçtüğünü gösterir.

AVE değeri genellikle faktör yüklerinin karesinin toplamına bölünerek hesaplanır.

$$\boldsymbol{AVE} = \frac{\left(\sum_{k=1}^{K} \boldsymbol{l}_k^2\right)}{\boldsymbol{K}} \tag{2.45}$$

AVE değerinin 0.5 veya daha yüksek olması istenen bir kriter olarak kabul edilir[106]. Bu, faktör tarafından açıklanan varyansın, toplam varyansa oranla en az yarısının göstergeler tarafından açıklandığı anlamına gelir. Diğer taraftan Hair ve arkadaşları [107], bileşik güvenirlik değerinin 0.60'ın üzeri olduğu durumlarda AVE'nin 0.50'den küçük olmasının yapı geçerliği için sorun olmadığını ifade etmiştir. Hair ve ark. için ise AVE için alt sınır 0.40'tır.[78]

Fornell ve Larcker [106] AVE değerini, yapısal eşitlik modelinde bir faktörün ölçüm modelinin geçerliliğini değerlendirmek için kullanmışlardır. Chin ve ark. [108] ise KEKK-YEM modellerinde AVE değerinin kontrol edilmesi için faktör yüklerini ve kovaryans yapısını analiz etmişlerdir.

2.10.1.3 Ayrışma Geçerliği

Kısmi en küçük kareler yapısı eşitlik modellerinde ayrışma geçerliği, bir faktör ile diğer faktörler arasındaki ayırt edilebilirliğin ölçüsüdür. Bu, bir faktörün diğer faktörlerle ilişkilerinden bağımsız olarak varyansının ne kadarını açıkladığını gösterir.

Ayrışma geçerliliği 3 farklı kritere göre kontrol edilebilir. Bunlar;

- Çapraz Yükler

- Fornell–Larcker Ölçütü
- HTMT Katsayısı

KEKK-YEM modellerinde çapraz yükler, farklı gizil değişkenler arasındaki ilişkileri ölçmek için kullanılan bir kavramdır. Çapraz yükler, bir gizil değişkenin göstergeler tarafından açıklanan varyansının başka bir gizil değişkenle olan ilişkisini ifade eder. Bu kurala göre, bir göstergenin, ilgili değişkendeki gösterge yükü, diğer değişkenlerdeki gösterge yüklerinden daha fazla olmalıdır[58, 109].

Diğer bir kriter Fornell-Larcker ölçütü, bir gösterge değişkenin kendi gizil değişkeniyle olan ilişkisini diğer gizil değişkenlerle olan ilişkileriyle karşılaştırır. Bu ölçüt, gösterge değişkenlerin ağırlıklı faktör yüklerini ve değişkenler arasındaki kovaryansları kullanarak hesaplanır. Eğer bir gösterge değişkenin kendi gizil değişkeniyle olan ilişkisi, diğer gizil değişkenlerle olan ilişkilerden daha büyük ise, Fornell-Larcker ölçütü sağlanmış olur. Bu durum, gösterge değişkenin daha fazla özgün varyansı açıkladığını ve ölçtüğü gizil değişkenle daha yüksek bir ölçüm ilişkisine sahip olduğunu gösterir. Fornell-Larcker ölçütü, ölçek geçerliliğini değerlendirmek ve gösterge değişkenlerin ayırt edici özelliklerini belirlemek için yaygın olarak kullanılan bir yöntemdir[106].

HTMT (Heterotrait-Monotrait) katsayısı ise, KEKK-YEM modellerinde çapraz yüklerin daha objektif bir şekilde değerlendirilmesi için kullanılan bir ölçüdür. HTMT katsayısı, iki gizil değişken arasındaki çapraz yüklerin, aynı gizil değişken arasındaki yüklerle karşılaştırıldığında ne kadar farklı olduğunu gösterir. Bu katsayı, 0 ile 1 arasında bir değer alır ve 0.90'dan küçük olması arzu edilen bir değerdir[86]. Daha düşük HTMT katsayısı, çapraz yüklerin daha iyi ayrılabilir olduğunu gösterir.

Çapraz yüklerin ve bu ölçütlerin kontrol edilmesi, KEKK-YEM modellerinin geçerliliğini ve ayrışma geçerliliğini değerlendirmek için önemlidir. Literatürde çapraz yüklerin ve bu ölçütlerin kontrol edilmesi için farklı yaklaşımlar ve formüller mevcuttur.

Örneğin, Fornell ve Larcker [106] çapraz yüklerin değerlendirilmesi için Fornell-Larcker ölçütünü önermişlerdir. Henseler vd. [86] ise HTMT katsayısının çapraz

yüklerin değerlendirilmesinde daha güvenilir bir alternatif olduğunu belirtmişlerdir.

2.10.2 Doğrusallık Varsayımı

KEKK-YEM modellerinde çoklu doğrusallık varsayımı, tahmin edici değişkenler arasında yüksek düzeyde çoklu doğrusallığın olmamasını ifade eder. Çoklu doğrusallık, tahmin edici değişkenlerin birbirleriyle yüksek derecede ilişkili olması durumunda ortaya çıkar ve tahmin edici değişkenler ile gizil değişkenler arasındaki ilişkilerin istikrarını bozabilir ve yorumlanmasını zorlaştırabilir[78].

Çoklu doğrusallığı kontrol etmek için genellikle varyans şişirme faktörü (VIF) değerleri incelenir. VIF, tahmin edilen regresyon katsayısının varyansının çoklu doğrusallık nedeniyle şişirilme derecesini ölçer. Genel olarak, 5 veya 10'dan büyük VIF değerleri çoklu doğrusallığın varlığını işaret eder ve sonuçların yorumlanmasında dikkatli olunması gerektiğini gösterir[78].

KEKK-YEM'nin geleneksel regresyon analizi gibi diğer yöntemlere göre çoklu doğrusallığa daha az duyarlı olduğu göz önünde bulundurulmalıdır. Bununla birlikte, sonuçların sağlamlığını sağlamak için çoklu doğrusallığın değerlendirilmesi önerilen bir uygulamadır[78].

2.10.3 Uyum İyiliği

YEM modellerinde, Ki-Kare (Chi-square - χ^2), NFI (Normed Fit Index), d_ULS (Degree of Unweighted Least Squares), ve d_G (Degree of Geodesic) gibi ölçütler, modelin uyum iyiliğini değerlendirmek için kullanılan önemli ölçütler arasındadır.

SRMR, modelin gözlemlenen veriler ile tahmin edilen değerler arasındaki uyumu ölçer. Düşük SRMR değerleri, daha iyi bir uyum iyiliği gösterir. SRMR değerinin 0.08 veya daha düşük olması kabul edilebilir bir uyum iyiliğini gösterir (Hu ve Bentler, 1999).

Ki-kare (χ^2), modelin uyum iyiliğini ölçmek için kullanılan yaygın bir ölçüttür. Ancak, Ki-kare değeri tek başına bir modelin uyum iyiliğini değerlendirmek için yeterli değildir. Çünkü Ki-kare değeri, örnekleme büyüklüğüne ve modelin

karmaşıklığına bağlı olarak etkilenebilir. Bu nedenle, Ki-kare değerinin yanında diğer uyum iyiliği ölçütlerinin de dikkate alınması önerilmektedir[63].

NFI (Normed Fit Index), modelin gözlemlenen verilere uyumunu ölçer. Yüksek NFI değerleri, daha iyi bir uyum iyiliği gösterir. NFI değeri genellikle 0.90 veya daha yüksek olmalıdır[110].

d_ULS (Degree of Unweighted Least Squares), modelin uyum iyiliğini ölçen bir ölçüttür. Daha düşük d_ULS değerleri, daha iyi bir uyum iyiliği gösterir. d_ULS değeri genellikle 0.08 veya daha düşük olmalıdır[86]

d_G (Degree of Geodesic), modelin uyum iyiliğini değerlendiren bir ölçüttür. Daha düşük d_G değerleri, daha iyi bir uyum iyiliği gösterir. d_G değeri genellikle 1.20 veya daha düşük olmalıdır[86]

Uyum iyiliğini değerlendirmek için tek bir ölçüt yerine birden fazla ölçütün birlikte değerlendirilmesi önerilmektedir[78].

Ayrıca Olya ve Henseler'e göre KEKK-YEM modellerinde uyum iyiliği değerlerinin sağlanması gibi bir zorunluluk bulunmamaktadır[79, 111]. Bu durum, uyum iyiliği şartı olan K-YEM yaklaşımına karşılık olarak KEKK-YEM yaklaşımını tercih etme sebebi olabilmektedir.

2.10.4 İçsel İlişkilerin Anlamlılığının Değerlendirilmesi

KEKK-YEM modellerinde içsel ilişkilerin anlamlılığını değerlendirmek için çeşitli istatistiksel yöntemler kullanılır. Bu yöntemler arasında Bootstrapping (Yeniden Örnekleme), Blindfolding (Tahmin Gücü Analizi), hipotez testleri, R^2 ve f^2 katsayılarının değerlendirilmesi sayılabilir.

2.10.4.1 Yeniden Örnekleme

KEKK-YEM modellerinde yeniden örnekleme, parametrik olmayan bir yöntemdir ve istatistiksel çıkarımların güvenilirliğini değerlendirmek için yaygın olarak kullanılmaktadır[78]. Bu yöntem, örneklem verilerinden tekrar tekrar örnekleme yaparak yeni örneklem dağılımları oluşturur ve bu dağılımlar üzerinden parametre tahminlerinin dağılımını hesaplar. Özellikle küçük örneklemlerle çalışırken parametre tahminlerinin yanı sıra standart hatalar, güven aralıkları ve

hipotez testlerinin güvenilirliğini sağlamak için yaygın olarak kullanılan bir yöntemdir.

Yeniden örnekleme, KEKK-YEM modellerinde modelin güçlüğünü değerlendirmek için de kullanılır. Orijinal veri setinin yapısını koruyarak birden çok örnekleme oluşturulur ve her bir örnekleme üzerinde modelin performansı ölçülür. Bu sayede, modelin kararlılığı, katsayıların istatistiksel anlamlılığı ve hedef değişkenin tahmin gücü gibi önemli değerlendirme kriterleri elde edilir[78].

Yeniden örnekleme ayrıca, KEKK-YEM modellerinde değişkenler arasındaki ilişkilerin anlamlılığını değerlendirmek için kullanılan çapraz geçerlilik analizlerinde de önemli bir rol oynar. Örneğin, çapraz geçerlilik analizlerinde, yeniden örnekleme kullanılarak yapısal yol katsayılarının istatistiksel anlamlılığı ve güven aralıkları hesaplanır, böylece değişkenler arasındaki ilişkilerin güçlüğü ve istatistiksel önemi değerlendirilebilir[86].

2.10.4.2 Tahmin Gücü Analizi

Bu analiz, her bir gözlemi tahmin etmek için modelin geri kalan gözlemleri kullanarak gerçek ve tahmin edilen değerler arasındaki farkları hesaplar [112].

Tahmin gücü analizi, modelin tahmin yeteneğini değerlendirmek için önemli bir araçtır. Tahmin edilen değerlerin gerçek değerlere ne kadar yakın olduğunu gösteren bir dizi istatistiksel ölçüt elde edilir. Bu ölçütler arasında kök ortalama kare hatası (root mean square error, RMSE), tahmin doğruluğu (prediction accuracy), R^2 değeri ve korelasyon katsayısı gibi değerler yer alır[112].

Bu analiz, KEKK-YEM modellerinde modelin öngörü yeteneğini değerlendirmek ve modelin geçerliliğini ve güvenilirliğini test etmek için yaygın olarak kullanılan bir yöntemdir. Tahmin edilen değerlerin gerçek değerlere ne kadar yakın olduğunu ölçerek modelin tahmin performansını değerlendirmek mümkün olur. Yeniden örnekleme ve tahmin gücü analizi (Q2) hesaplamaları, KEKK-YEM yazılımları aracılığıyla otomatik olarak yapılabilir. Bu hesaplamalar, modelin güvenilirliği, istatistiksel anlamlılığı ve tahmin yeteneği hakkında önemli bilgiler sağlar.

2.10.4.3 Hipotez Testleri

Bu yöntemde genellikle parametre tahminlerinin sıfır hipotezi karşısında anlamlılığının test edilmesi kullanılır. Sıfır hipotez, ilişkinin var olmadığını veya parametrenin sıfıra eşit olduğunu ifade eder. İlgili parametre tahmininin anlamlılığı, hipotez testi sonucunda elde edilen p değeri ile değerlendirilir. Eğer p değeri belirlenen anlamlılık düzeyinden küçükse (genellikle $p<0.05$), ilişki anlamlı olarak kabul edilir.

2.10.4.4 R^2 ve f^2 Katsayıları

Bu değerlendirmelerin yanı sıra, içsel ilişkilerin etki büyüklüğü de değerlendirilebilir. Özellikle KEKK-YEM modellerinde, standartlaştırılmış etki büyüklüğü (standardized effect size) ölçütü kullanılarak, ilişkilerin gücü ve önemi hakkında bilgi elde edilebilir.

R^2, bağımlı değişkenin varyansının bağımsız değişkenler tarafından ne kadar açıklandığını gösterir. Bu katsayı, modelin açıklama gücünü ölçmek için kullanılır. R^2 değeri 0 ile 1 arasında değişir ve ne kadar yüksekse, bağımlı değişkenin bağımsız değişkenler tarafından o kadar iyi açıklandığı anlamına gelir. Genellikle R^2 değeri yüksek olan modeller, daha iyi uyum sağlayan ve daha iyi tahminler yapan modeller olarak kabul edilir[87]. B endojen ve tahminci değişken arasındaki yol katsayısı, cor ise endojen ve tahminci değişken arasındaki korelasyon değeri olmak üzere, R^2 değeri şu formülle hesaplanır.

$$\boldsymbol{R^2 = \sum_{j}^{n} \widehat{\beta_j}\ cor(y, x_j)} \tag{2.46}$$

Diğer yandan, R^2 değeri modelin karmaşıklığına bağlı olarak yanıltıcı olabilir. Özellikle çok sayıda bağımsız değişkenin olduğu karmaşık modellerde, R^2 değeri yüksek çıkabilir ve bu durum modelin gerçek açıklama gücünü aşırı tahmin edebilir[87].

Düzeltilmiş R^2, R^2 değerini serbestlik derecesi ve örneklem büyüklüğü gibi faktörlerle düzeltir. Bu düzeltme, modelin karmaşıklığına bağlı olarak düzeltilmiş R^2 değerinin daha gerçekçi bir tahmin sağlamasını sağlar. Düzeltilmiş R^2 değeri,

her bir bağımsız değişkenin ek açıklama gücünü dikkate alarak daha dengeli bir değerlendirme yapmayı sağlar[87].

n örneklem büyüklüğü, k modeldeki tahminci yapıların sayısı olmak üzere düzeltilmiş R^2 şu formülle hesaplanır.

$$\mathbf{R^2_{adj} = 1 - (1 - R^2) . \frac{n - 1}{n - k - 1}} \quad (2.47)$$

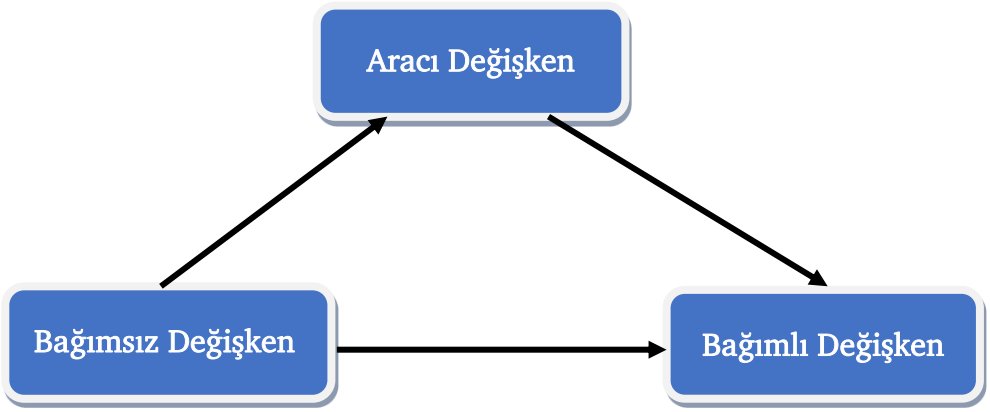
f^2 ise değişkenler arasındaki ilişkinin etki büyüklüğünü gösterir. Bu katsayı, bağımsız değişkenlerin bağımlı değişkendeki değişikliği ne ölçüde açıkladığını ifade eder. f^2 değeri 0 ile ∞ arasında değişir. Genellikle f^2 değeri 0.02, 0.15 ve 0.35 olmak üzere üç farklı etki büyüklüğü sınıfına ayrılır. Değerlerin artmasıyla etki büyüklüğü de artar[113].

f^2 değeri, ilgili değişkenin modele dahil edildiği ve hariç tutulduğu durumlardaki R^2 değerlerine göre şu formülle hesaplanır.

$$\mathbf{f^2 = \frac{\left(R^2_{dahil} - R^2_{hariç}\right)}{\left(1 - R^2_{dahil}\right)}} \quad (2.48)$$

2.11 Aracılık Analizi

Aracılık analizi, iki değişken arasındaki ilişkiyi açıklayan temel mekanizmaları incelemek için sosyal ve davranışsal bilimlerde sıklıkla kullanılan bir istatistiksel yöntemdir[114]. Bağımsız bir değişken ile bağımlı bir değişken arasındaki ilişkinin kısmen veya tamamen bir veya daha fazla ara değişken tarafından açıklanıp açıklanmadığını test etmek, yani aracı (mediatör) olarak adlandırılan bir veya daha fazla ara değişken tarafından açıklanıp açıklanmadığını test etmek, aracılık analizinin temel amacıdır[115].



Şekil 2.8 Aracılık modelinin kavramsal diyagramı

Aracılık analizi, müdahalenin, tedavinin veya manipülasyonun sonuç değişkeni üzerindeki etkisinin temel mekanizmasını belirlemek için sıklıkla kullanılır[116]. Bu teknik, psikoloji[117], tıp[115], eğitim[114], ekonomi[118] ve halk sağlığı[119] gibi çeşitli alanlarda uygulanmıştır.

Aracılık analizi için farklı yöntemler mevcuttur. Bunlar arasında nedensel adımlar yaklaşımı, katsayıların çarpımı yaklaşımı ve yapısal eşitlik modeli yaklaşımı yer almaktadır[120]. Nedensel adımlar yaklaşımı, bağımsız değişkenin medyatör aracılığıyla bağımlı değişkendeki doğrudan ve dolaylı etkilerini incelemek için bir dizi regresyon modeli test etmeyi içerir. Katsayıların çarpımı yaklaşımı ise, bağımsız değişken ile medyatör arasındaki yol ve medyatör ile bağımlı değişken arasındaki yolun katsayılarını çarpmayı içerir. Yapısal eşitlik modeli yaklaşımı ise, bağımsız değişkeni, aracı değişkeni ve bağımlı değişkeni içeren bir yol modeli test etmeyi içerir.

Aracılık analizi aynı zamanda kısmi aracılık ve tam aracılık gibi farklı aracılık türlerini de içermektedir[116]. Kısmi aracılık, bağımsız değişken ve bağımlı değişken arasındaki ilişkinin bir kısmını, ancak tamamını değil, açıklayan aracının varlığına işaret eder. Tam aracılık ise bağımsız değişken ve bağımlı değişken arasındaki ilişkinin tamamını açıklayan aracının varlığı anlamına gelmektedir.

Zamansal öncelik problemi, ölçüm problemi ve potansiyel karıştırıcı değişkenler gibi farklı sorunlar aracılık analizi yaparken dikkate alınması gereken diğer konulardandır[115]. Bu konular sonuçların geçerliliği ve bulguların yorumlanmasını etkileyebilir.

Tüm bu nedenlerden dolayı, aracılık analizi, bilimsel araştırmalarda değişkenler arasındaki altta yatan mekanizmaları incelemek için değerli bir araçtır.

2.11.1 Sobel Testi

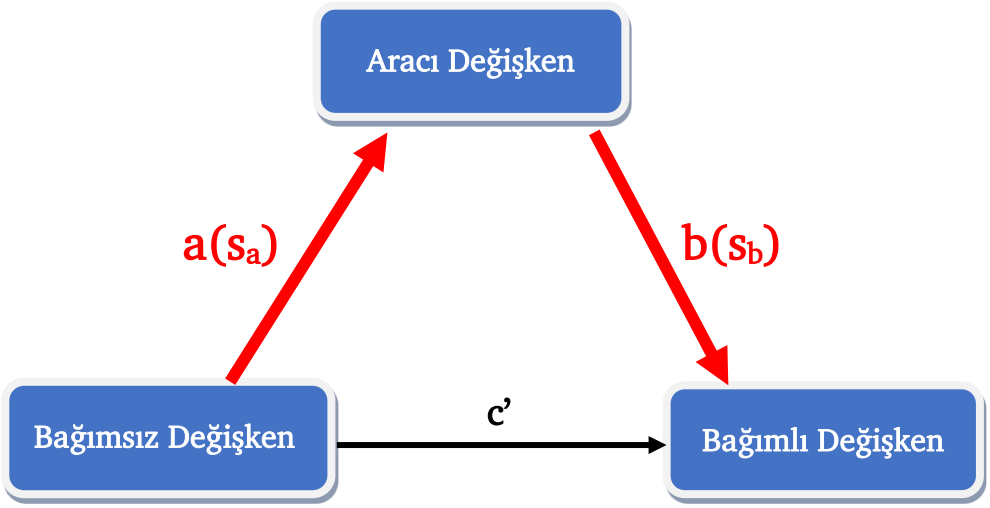
Sobel testi, aracı değişkenlerin etkilerinin anlamlılığını test etmek için kullanılan bir istatistiksel yöntemdir. Bu test, bir bağımsız değişkenin etkisinin, bir aracı değişken üzerinden bağımlı değişkene olan dolaylı etkisini değerlendirmek amacıyla kullanılır[121].

Sobel testinin geliştirilen 2 farklı versiyonu daha mevcuttur. Bunlar Aroian ve Goodman testleridir. Sobel testi, bağımlı değişkene olan dolaylı etkinin anlamlılığı için en sık test edilen yöntemdir, ancak bazı sınırlamaları vardır. İlk olarak, çok güçlü değildir, yani bir etkinin gerçek olduğu durumlarda bile aracılık etkisini tespit etme şansı düşüktür[122]. İkinci olarak, Sobel testi örneklemin büyüklüğüne duyarlıdır, yani büyük örneklerde küçük örneklerden daha çok aracılık etkisi tespit etme eğilimindedir[122].

Aroian ve Goodman testleri, Sobel testinden daha güçlüdür, ancak aynı zamanda daha karmaşıktır. Aroian testi, IV, aracı değişken ve DV arasındaki yol katsayılarının çarpımına dayanırken, Goodman testi, toplam etki ile doğrudan etki arasındaki farka dayanır Aroian testi parametrik olmayan bir testken, Goodman testi parametrik bir testtir. [120, 122]

Aracılık için hangi testin kullanılacağı mevcut örneklem büyüklüğüne bağlıdır. Eğer güç önemli değilse, Sobel testi iyi bir seçenektir. Güç önemliyse, Aroian veya Goodman testleri daha iyi seçenekler olabilir.

Sobel testi, aracılık etkisinin anlamlılığını değerlendirmek için iki adımdan oluşur. İlk adımda, aracı değişkenin bağımsız değişken üzerindeki etkisi doğrudan bir regresyon modeli kullanılarak hesaplanır. İkinci adımda ise, aracı değişkenin bağımsız değişken üzerindeki etkisi dolaylı etki yoluyla hesaplanır. Dolaylı etki yolu, aracı değişkenin bağımlı değişken üzerindeki etkisini aracı bir değişken üzerinden aktarır. Sobel testi, doğrudan ve dolaylı etki yollarının standart hatalarını ve katsayılarını kullanarak aracı etkisinin anlamlılığını test eder[121].



Şekil 2.9 Sobel test (aracı değişken illüstrasyonu)

Şekil 2.9'da gösterilen sembollerin anlamları sırasıyla şöyledir.

a = Bağımsız değişken ve aracı değişken arasındaki ilişki için ham (standartlaştırılmamış) regresyon katsayısı

s_a = a'nın standart hatası.

b = Aracı ile bağımlı değişken arasındaki ilişki için ham katsayı (bağımsız değişken aynı zamanda bağımlı değişkenin bir göstergesi olduğunda).

s_b = b'nin standart hatası.

c' =Bağımsız değişkenin, bağımlı değişken üzerindeki direkt etkiyi ifade eder.

Aracılık modeli doğru bir şekilde belirtilirse, c, a, b ve c' yolları, bazen sıradan en küçük kareler veya çoklu regresyon ile tahmin edilebilir. Bazı durumlarda, çoklu regresyon yerine diğer tahmin yöntemleri (örneğin, lojistik regresyon, çok düzeyli modelleme ve yapısal eşit modelleme) kullanılmalıdır. Hangi veri analizi yöntemi kullanılırsa kullanılsın, arabuluculuğu test etmek için gerekli adımlar aynıdır.

Öncelikle bağımsız değişkenin aracı değişkeni tahminlediği bir regresyon analizi çalıştırılır. Böylece a ve s_a değerleri hesaplanır. Bağımsız ve aracı değişkenin birlikte bağımlı değişkeni tahminlediği ikinci bir regresyon analizi daha çalıştırılır. Bu adım ise b ve s_b değerlerini verecektir. Bu bölümde hesaplanan standart hata değerlerinin hiçbir zaman negatif olamayacağına dikkat edilmelidir.

Hesaplanan bu 4 değer kullanılarak, aracılık testi için kritik değer olan z değeri hesaplanır. Sırasıyla Sobel, Aroian ve Goodman test eşitlikleri şu şekildedir.

$$z = \frac{a.b}{\sqrt{b^2.{s_a}^2 + a^2.{s_b}^2}} \tag{2.49}$$

$$z = \frac{a.b}{\sqrt{b^2.{s_a}^2 + a^2.{s_b}^2 + {s_a}^2.{s_b}^2}} \tag{2.50}$$

$$z = \frac{a.b}{\sqrt{b^2.{s_a}^2 + a^2.{s_b}^2 - {s_a}^2.{s_b}^2}} \tag{2.51}$$

Bu değerin anlamlılığı, normal dağılım varsayımı altında Z-testi kullanılarak test edilir. Eğer hesaplanan Z değeri, kritik Z değerinden büyükse aracı etkisi anlamlı kabul edilir[121].

Bu testin varsayımları şu şekilde listelenebilir.

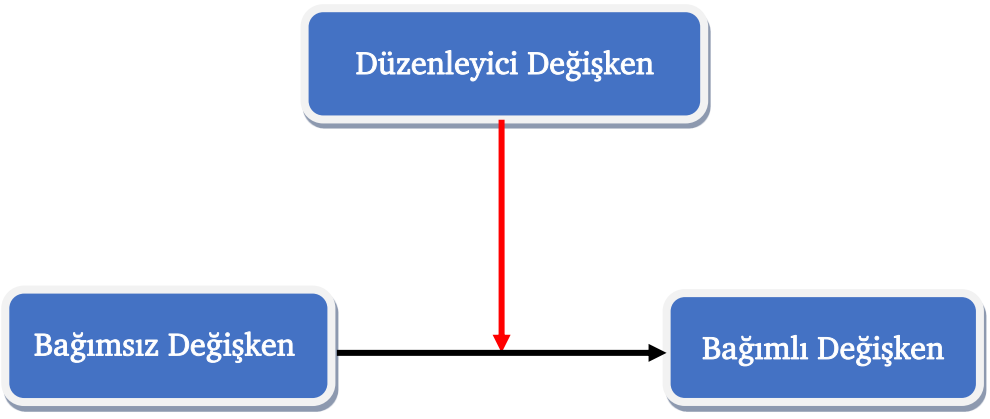
- Bağımsız değişken, aracı değişken ve bağımlı değişken arasında doğrusal bir ilişki bulunmalıdır: Sobel testi, doğrusal ilişkilerin olduğu durumlarda uygulanır. Bu varsayım, değişkenlerin doğrusal bir ilişkiye sahip olduğunu ve normal dağıldığını ifade eder.
- Bağımsız değişkenin etkisi, aracı değişken tarafından tam olarak aracılık edilmelidir: Sobel testi, bağımsız değişkenin etkisinin tamamen aracı değişken üzerinden geçtiği durumları değerlendirir. Bu varsayım, aracı değişkenin bağımsız değişken ile bağımlı değişken arasındaki ilişkide gerçekten bir aracılık rolü oynadığını gösterir. Eğer aracı değişkenin etkisi zayıf veya yoksa, Sobel testi uygun sonuçlar vermeyebilir.
- Bağımsız değişken, aracı değişken ve bağımlı değişken arasında normal dağılım sağlanmalıdır: Sobel testi, Z-testi kullanarak aracılık etkisinin anlamlılığını test eder. Bu nedenle, testin doğru sonuçlar vermesi için değişkenlerin normal dağılıma sahip olması önemlidir. Normal dağılım varsayımı ihlal edildiğinde, alternatif yöntemlerin kullanılması gerekebilir[121].

Sobel testinin bu varsayımları dikkate alınmalı ve testin sonuçları yorumlanırken göz önünde bulundurulmalıdır. Varsayımların sağlanmadığı durumlarda, alternatif yöntemlerin kullanılması veya daha geniş bir analiz yapılması gerekebilir.

Ayrıca sobel testi, sadece tek bir düzenleyici değişkenin etkisini değerlendirmek için kullanılır. Birden fazla düzenleyici değişken varsa, alternatif yaklaşımlar kullanılması gerekir.

2.12 Düzenleyici Değişken Analizi

Düzenleyici değişken analizi, sosyal ve davranışsal bilimlerde, iki değişken arasındaki ilişkinin üçüncü bir değişken olan düzenleyicinin seviyelerine bağlı olarak nasıl değiştiğini araştırmak için kullanılan bir istatistiksel tekniktir[114]. Hayes'e göre, düzenleyici değişken analizi araştırmacılara öngördükleri ve sonuç değişkeni arasındaki ilişkinin farklı düzenleyici (moderatör, ılımlaştırıcı) değişkeni seviyelerinde nasıl değiştiğini inceleyerek yardımcı olur[116]. Teknik, bir ilişkinin hangi sınırlar altında geçerli olduğunu keşfetmek için yaygın olarak kullanılır ve değişkenler arasındaki ilişkiyi değiştiren koşullara ilişkin bilgiler sağlayabilir[123].



Şekil 2.10 Düzenleyici değişken modelinin kavramsal diyagramı

Görsel yaklaşım, regresyon temelli yaklaşım ve varyans analizi (ANOVA) temelli yaklaşım olmak üzere düzenleyici değişken analizi yapmak için farklı yöntemler bulunmaktadır[124]. Görsel yaklaşım, düzenleyici değişkenini X-ekseninde yerleştirerek öngörücü ve sonuç değişkenlerinin bir dağılım grafiği oluşturulmasını içerir. Regresyon temelli yaklaşım, bir etkileşim teriminin

regresyon modeline eklenmesini gerektirir ve düzenleyici değişkeninin, öngörücü ve sonuç değişkenleri arasındaki eğimi değiştirmesine izin verir. ANOVA temelli yaklaşım, etkileşim teriminin ANOVA modelindeki anlamlılığını test etmeyi içerir.

Düzenleyici değişken analizi ayrıca basit moderasyon, modere moderasyon ve aracılı moderasyon gibi farklı moderasyon türlerine sahiptir[116]. Basit moderasyon, öngörücü ve sonuç değişkeni arasındaki ilişkinin gücünün veya yönünün düzenleyici değişkeni seviyelerine bağlı olarak değiştiği durumlarda gerçekleşir. Modere moderasyon, düzenleyicilik etkisinin gücünün veya yönünün dördüncü bir değişkenin seviyelerine bağlı olarak değiştiği durumlarda ortaya çıkar. Aracılı moderasyon, öngörücü ve sonuç değişkeni arasındaki ilişkinin aracının düzenleyici tarafından etkilendiği durumlarda meydana gelir.

Düzenleyici değişken analizi yapılırken düzenleyicinin ölçümü, düzenleyici etkilerinin analizi için uygun yöntemlerin seçimi ve karıştırıcı değişkenlerin varlığı gibi birçok konu dikkate alınmalıdır[123]. Bu konular sonuçların geçerliliğini ve bulguların yorumlanmasını etkileyebilir.

Genel olarak, düzenleyici değişken analizi, bir ilişkinin hangi koşullar altında var olduğunu araştırmak ve değişkenler arasındaki ilişkiyi ılımlayan koşullar hakkında bilgi sağlamak için özellikle sosyal ve davranışsal araştırmalarda kullanılan değerli bir istatistiksel tekniktir.

Düzenleyici değişken modellerinin süreçleri, düzenleyici değişkenin nümerik ve kategorik olmasına bağlı olarak iki farklı şekilde gerçekleşir.

2.12.1 Kategorik Değişkenlerin Düzenleyici Etkilerini Modelleme

Kategorik değişkenlerin düzenleyici etkilerini KEKK-YEM'de modelleme, bağımsız ve bağımlı değişkenler arasındaki ilişkiyi kategorik bir düzenleyici değişkenin etkisiyle incelemeyi içerir. Bu yaklaşım, değişkenler arasındaki ilişkinin farklı gruplar veya kategoriler arasında nasıl değişebileceğini daha derinlemesine anlaşılmasını sağlar.

KEKK-YEM'de kategorik düzenleyici etkilerini modellemek için yaygın olarak kullanılan bir yöntem, grup yaklaşımıdır. Bu yaklaşım, düzenleyici değişkeninin her bir kategorisi için ayrı modeller oluşturmayı ve bu modeller arasında tahmin

edilen yol katsayılarını veya etki büyüklüklerini karşılaştırmayı içerir ve böylece değişkenler arasındaki ilişkinin gruplar arasında nasıl farklılaştığını anlaşılmasına yardımcı olur.

Diğer bir yaklaşım ise Kısmi En Küçük Kareler Çoklu Grup Analizi / Partial Least Squares Multi-Group Analysis (KEKK-ÇGA/PLS-MGA) dir. Burada düzenleyici değişkeninin farklı kategorileri aynı anda analiz edilerek model parametreleri arasındaki anlamlı farklar test edilir. Bu yöntem, farklı gruplar arasındaki yol katsayılarını veya etki büyüklüklerini doğrudan karşılaştırma imkânı sağlar ve düzenleyici etkinin istatistiksel kanıtlarını sunar.

KEKK-ÇGA, yapısal denklem modellemesi çerçevesinde gruplar arasındaki farklılıkların incelenmesine olanak tanıyan kısmi en küçük kareler yaklaşımının bir uzantısıdır[86]. KEKK-ÇGA, özellikle farklı karakteristiklere sahip olan veya bir nüfusun farklı segmentlerini karşılaştırmak istendiği durumlarda kullanışlıdır. Bu yöntem, gruplar arasındaki farklılıkların öneminin değerlendirilmesine ve ilişkilerin gruplar arasındaki değişmezliği incelemeye olanak sağlar.

KEKK-ÇGA'da, her bir grup için grup özelinde modeller ayrı ayrı tahmin edilir ve her bir alt grubun özgün özelliklerini yakalar[125]. Analiz, her bir grup için yapısal modelin ayrı ayrı tahmin edilmesini içeren bir dizi iteratif adımdan oluşur, ardından tahmin edilen parametrelerin gruplar arasındaki farklarının anlamlılığı karşılaştırılır. Grup karşılaştırmaları, permütasyon testleri veya yeniden örnekleme yöntemleri kullanılarak parametre farklılıklarının anlamlılığına bakılarak yapılır[102]

KEKK-ÇGA'nın temel amacı, gizil değişkenler arasındaki ilişkilerin gruplar arasında nasıl farklılaştığını anlamak ve heterojenliğin potansiyel kaynaklarını belirlemektir. Grup özelinde modelleri dikkate alarak, araştırmacılar grup özelindeki varyasyonları hesaba katabilir ve bulgularının daha doğru ve güvenilir yorumlarını yapabilirler[78].

KEKK-YEM'de kategorik düzenleyici etkilerini modellemekle ilgili kısıtlamalardan biri, grup değişmezliği varsayımıdır, bu da değişkenler arasındaki ilişkilerin tüm gruplar için tutarlı olması anlamına gelir. Bu varsayımın ihlali, yanlı veya yanlış sonuçlara yol açabilir. Düzenleyici değişken analizinin geçerliliğini sağlamak için

gruplar arasında ölçüm değişmezliğinin ve yapısal değişmezliği değerlendirilmesi önemlidir.

Ayrıca, düzenleyici değişkenindeki kategorilerin sayısı dikkate alınmalıdır. Az sayıda kategoriye sahip olmak, anlamlı düzenleyici etkilerini tespit etme gücünü sınırlayabilir, çok sayıda kategori ise küçük grup boyutlarına ve güvenilir olmayan tahminlere neden olabilir. Araştırmacılar, yeterli sayıda grup ve her bir grup içinde yeterli örneklem büyüklüğünü koruma arasında bir denge sağlamaya çalışmalıdır.

2.12.2 Nümerik Değişkenlerin Düzenleyici Etkilerini Modelleme

KEKK-YEM'de sürekli değişkenlerin düzenleyici etkilerini modellemek, iki yapı arasındaki ilişkiyi etkileyen sürekli değişkenlerin farklı düzeylerindeki etkisini incelemek için kullanılan değerli bir tekniktir. Nümerik düzenleyici etkileri, bağımsız değişken ile bağımlı değişken arasındaki ilişkinin düzenleyici değişkenin farklı düzeylerinde değiştiği durumu ifade eder[78].

KEKK-YEM'de nümerik düzenleyici etkileri modellemek için yapısal modelde bağımsız değişken, düzenleyici değişken ve bu değişkenlerin çarpımı olan etkileşim terimleri kullanılır. Etkileşim terimi, bağımsız değişken ve düzenleyici değişkenin çarpımının bağımlı değişken üzerindeki çarpan etkisini yakalar ve ilişkinin gücünün veya yönünün düzenleyici değişkenin farklı düzeylerinde nasıl değiştiğini gösterir[78, 126].

Düzenleyici etkinin anlamlılığını değerlendirmek için araştırmacılar, etkileşim terimi için güven aralıklarını ve p-değerlerini elde etmek için yeniden örnekleme yöntemlerini kullanabilir. Yeniden örnekleme, verilerin birden fazla kez yeniden örnekleme yaparak etkileşim etkisi tahminlerinin dağılımını oluşturmayı sağlar ve sağlam istatistiksel çıkarımlar yapmayı mümkün kılar[78].

Ayrıca, düzenleyici etkinin görselleştirilmesi ilişkiyi daha iyi anlamak için önemlidir. Etkileşim etkisinin çizelgelenmesi, düzenleyici değişkenin farklı düzeylerinde etkinin yönü ve büyüklüğü hakkında bilgi sağlayabilir[87].

Basit Eğim Analizi (Simple Slope Plot), Kısmi En Küçük Kareler Yapısal Eşitlik Modellemesinde bir nümerik değişkenin iki nümerik değişken arasındaki etkileşim etkisini incelemek ve yorumlamak için kullanılan bir görsel araçtır. Bu grafik,

bağımlı değişken ile bağımsız değişken arasındaki ilişkinin farklı düzenleyici seviyelerindeki gücünü ve yönünü görsel olarak temsil eder.

Basit Eğim Analizi, düzenleyici değişkeninin her bir seviyesi için ayrı ayrı regresyon çizgilerinin tahmin edilmesiyle oluşturulur[127]. Bu, bağımlı değişkenin düzenleyicinin farklı seviyelerindeki koşullu etkilerini incelemek için araştırmacılara olanak sağlar. Regresyon çizgilerinin eğimleri ilişkinin gücünü temsil ederken, yönleri düzenleyici etkinin doğasını gösterir.

KEKK-YEM'de Basit Eğim Analizi oluşturmak için araştırmacılar genellikle Hayes'in açıkladığı yöntemi takip eder[116]. İlk olarak, veriler düzenleyici değişkeninin 3 farklı seviyesine göre bölünür. Bu bölme işleminde aritmetik ortalama veya ortanca değerler tercih edilebilir. Ortalama kullanılırsa düzenleyici değişkenin aritmetik ortalaması hesaplanır ve bu değerden bir standart sapma çıkarılarak ve bir standart sapma eklenerek 3 farklı seviye elde edilir. Eğer ortanca kullanılırsa, düzenleyici değişkenin 1. kartil, 2. kartil (ortanca, medyan) ve 3. kartil değerleri hesaplanır. Ardından, her bir seviye için ayrı ayrı regresyon modelleri tahmin edilir ve regresyon katsayıları ile standart hataları elde edilir. Bu katsayılar, düzenleyici değişkenini x-ekseninde ve bağımlı değişkeni y-ekseninde görselleştirmek için kullanılır.

Simple Slope plotu yorumlamak, regresyon çizgilerinin eğimlerini ve güven aralıklarını incelemeyi içerir. Eğimlerin sıfırdan anlamlı olarak farklı olması, bağımsız ve bağımlı değişken arasındaki ilişkinin düzenleyici seviyelerine göre değiştiğini gösterir. Araştırmacılar ayrıca düzenleyicilerin farklı kategorileri arasındaki eğimleri karşılaştırarak düzenleyici etkisinin doğasını ve büyüklüğünü belirleyebilir.

SmartPLS ve WarpPLS gibi birçok yazılım paketi, KEKK-YEM analizlerinde Simple Slope Plot oluşturmak için araçlar ve işlevler sağlar. Bu araçlar, regresyon katsayılarının tahmin edilmesi ve çizgilerin çizilmesi sürecini otomatikleştirerek araştırmacıların düzenleyici etkilerini keşfetmelerini ve yorumlamalarını kolaylaştırır.



Şekil 2.11 Smartpls programında elde edilmiş bir simple slope plotun görünümü

2.13 PROCESS Macro

Process Macro (diğer adıyla Hayes Process), Andrew F. Hayes tarafından geliştirilen, sosyal bilimler alanında aracı değişken, düzenleyici değişken ve koşullu süreç analizlerini yapmak için kapsamlı bir metodolojik yaklaşımdır [127]. Araştırmacılara değişkenler arasındaki karmaşık ilişkileri incelemelerine ve bu ilişkilerin altında yatan mekanizmaları keşfetmelerine olanak sağlayan yaygın olarak kullanılan bir istatistiksel araçtır. SPSS veya SAS gibi istatistiksel analiz yazılımlarında kullanılabilir.

Process Macro, bağımsız değişkenlerin bir bağımlı değişkene etkilerini aracı değişkenler aracılığıyla analiz etmek için regresyon temelli bir yaklaşımı benimser[127]. Araştırmacılara bağımsız değişkenlerin sonuç değişkeni üzerindeki doğrudan ve dolaylı etkilerini değerlendirmelerine, aynı zamanda nedensel yol üzerindeki ara değişkenlerin aracılık rolünü incelemelerine olanak tanır. Bu noktada belirtilmelidir ki, Process Macro sadece nümerik olan aracı değişkenleri test edebilir. Kategorik aracı değişkenler için uygun değildir.

Ayrıca, Process Macro düzenleyici analizi için analitik yeteneklerini genişletir ve bağımsız değişken ile bağımlı değişken arasındaki ilişkinin üçüncü bir değişken

tarafından nasıl etkilendiğini araştırır[127]. Araştırmacılara ilişkinin gücünün veya yönünün düzenleyici değişkeninin farklı seviyeleri arasında farklı olup olmadığını değerlendirmelerine olanak sağlar.

Process Macro, yeniden örnekleme tekniğini kullanarak dolaylı ve düzenleyici etkilerin istatistiksel anlamlılığını tahmin eder[127].

2.14 KEKK-YEM Analizlerinde İncelenen Etkiler

KEKK-YEM modelleri, farklı etkileri analiz edebilir. Örneğin, araştırmacılar, değişkenler arasındaki nedensel etkileri ölçmek için yapısal eşitlik modelleri kullanabilirler. Bu, bir değişkenin diğer değişkenleri ne şekilde etkilediğini anlamak için kullanışlıdır.

2.14.1 Doğrudan Etki

Doğrudan etki, aracı değişkenler dikkate alınmadan bir değişkenin diğerine doğrudan etkisini temsil eder. Başka bir deyişle, araya giren değişkenlerin etkisini hariç tutarak iki değişken arasındaki ilişkinin gücünü ve yönünü gösterir[12]. Doğrudan etkiler, modeldeki değişkenleri birbirine bağlayan yollar boyunca parametreleri inceleyerek hesaplanır[128]. Bu etkiler, değişkenler arasındaki doğrudan ilişkileri anlamaya yardımcı olur ve doğrudan ve dolaylı etkileri içeren bir değişkenin genel etkisini hesaplamak için önemlidir[12].



Şekil 2.12 Doğrudan etkinin kavramsal diyagramı

Bu noktada iki değişken arasındaki etkiyi ifade eden değer, yol katsayısı şeklinde ifade edilir ve β ile sembolize edilir. Kısmi En Küçük Kareler Yapısal eşitlik modellemesinde B değeri, bağımlı ve ve bağımsız değişken arasındaki kovaryans ve bağımsız değişkenin standart sapması kullanılarak şu şekilde hesaplanır.

$$\boldsymbol{\beta} = \frac{\mathbf{cov}(\mathbf{X}, \mathbf{Y})}{\boldsymbol{\sigma X}} \tag{2.52}$$

Yol katsayılarının t-istatistiği ve p-değeri hesaplamaları için Yeniden örnekleme veya Jackknife yöntemleri kullanılabilir. Her iki yöntem de KEKK-YEM'de içsel modelin istatistiksel anlamlılığını değerlendirmek için yaygın olarak kullanılır.

Yeniden örnekleme yöntemi kullanılarak yol katsayılarının t-istatistiği ve p-değeri hesaplamaları şu adımlarla gerçekleştirilir:

- Yeniden örnekleme yapılır: Veri setinden, orijinal örnekleme büyüklüğüne eşit büyüklükte birden çok alt örneklem çekilir. Bu örneklemler, orijinal veri setinin rastgele tekrarlarını içerir.
- Her bir yeniden örnekleme için model yeniden tahmin edilir: Her yeniden örnekleme üzerinde KEKK-YEM modeli yeniden tahmin edilir ve yol katsayıları elde edilir.
- Yol katsayılarının dağılımı hesaplanır: Birden çok yeniden örnekleminde elde edilen yol katsayılarına dayanarak, bu katsayıların dağılımı hesaplanır. Buna göre, her bir yol katsayısı için ortalama ve standart sapma hesaplanır.
- Standart hata ve t-istatistiği hesaplanır: Yol katsayılarının standart hataları, ilgili yol katsayısının standart sapmasının karekökü olarak hesaplanır.

$$\boldsymbol{\sigma_M^{\beta}} = \sqrt{\boldsymbol{\sigma\beta}} \tag{2.53}$$

Ardından, t-istatistiği yol katsayısının tahmini ile standart hatanın bölümü olarak hesaplanır.

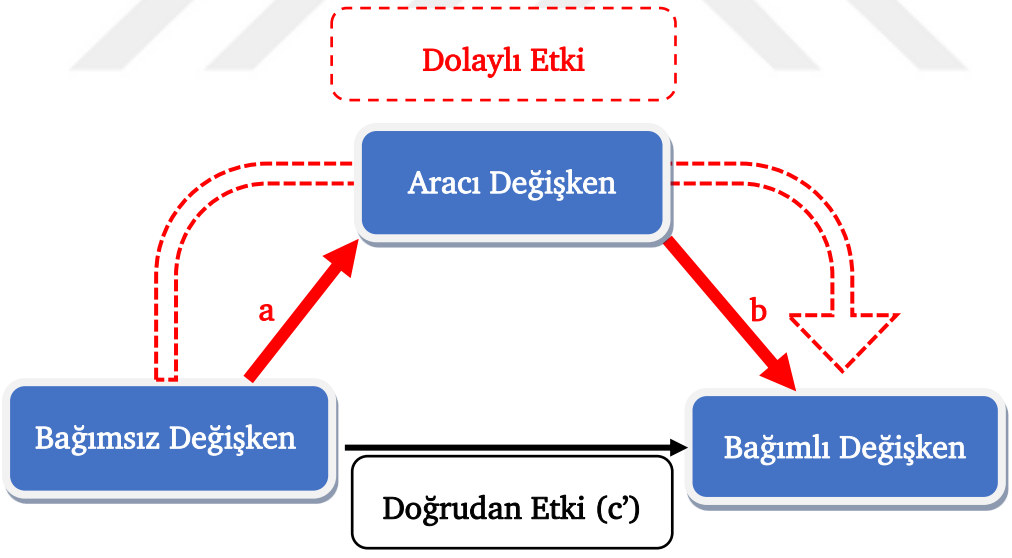
$$\boldsymbol{t_{\beta}} = \frac{\boldsymbol{\beta}}{\boldsymbol{\sigma_M^{\beta}}} \tag{2.54}$$

- p-değeri hesaplanır: Elde edilen t-istatistiği ve serbestlik derecesine (yeniden örnekleme sayısı eksi bir) dayanarak, p-değeri belirlenir. Bu, t-istatistiğin gözlenen veya daha büyük bir değeri elde etme olasılığını ifade eder.

Jackknife Yöntemi de yol katsayılarının t-istatistiği ve p-değeri hesaplamaları için kullanılabilen bir yöntemdir. Jackknife Yöntemi, veri setinden tekrarlı alt örneklem çekerek ve bu örneklemler üzerinde modeli yeniden tahmin ederek işleyişini gerçekleştirir.

2.14.2 Dolaylı Etki

Dolaylı etki, bir aracı değişken veya bir dizi aracı değişken aracılığıyla bir değişkenin diğeri üzerindeki etkisini ifade eder[129]. Başka bir deyişle, bir veya daha fazla araya giren değişken aracılığıyla bir değişkenin sonuç değişkeni üzerindeki etkisini temsil eder[128]. Dolaylı etkiler, bağımsız değişkenden aracı veya aracıları içeren bağımlı değişkene yönlendirilen yolların her biri boyunca parametrelerin çarpılması ve toplanmasıyla hesaplanır[128]. Dolaylı etkileri incelemek, değişkenler arasındaki ilişkilerde aracı değişkenlerin rolünü belirlemeye yardımcı olduğundan, değişkenlerin birbirini etkilediği temel mekanizmaları ve süreçleri anlamak için çok önemlidir[130].



Şekil 2.13 Aracılık etkisinin istatistiksel diyagramı

Dolaylı ilişki, bağımsız bir değişkenden bir aracıya yönelen ilişki katsayısı a ve aracıdan bağımlı değişkene yönelen ilişki katsayısı b olmak üzere a ve b değerlerinin çarpılmasıyla hesaplanır.

2.14.3 Toplam Etki

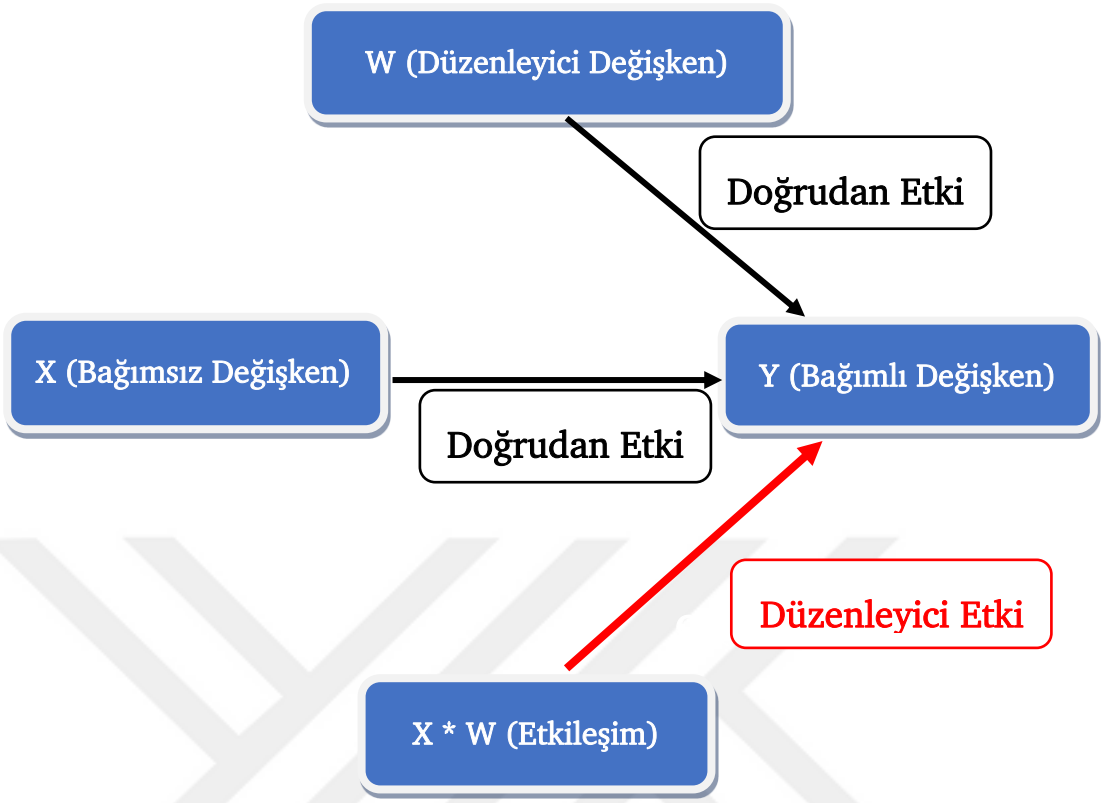
Toplam etki bir değişkenin diğerine hem doğrudan hem de dolaylı etkilerinin birleşik etkisini temsil eder. Bu, araya giren değişkenler aracılığıyla olanları da içeren bağımsız bir değişkenin bağımlı değişken üzerindeki genel etkisini yakalar. Toplam etkiyi hesaplamak için, doğrudan etki ve değişkenler arasındaki tüm dolaylı etkiler toplanmalıdır[128]. Toplam etkiyi anlamak, modeldeki değişkenler arasındaki tam ilişkiyi anlamak için önemlidir ve doğrudan ve dolaylı etkileşimlerin karmaşık bir şekilde nasıl etkileşime girdiğini çözmeye yardımcı olur[129]. Toplam etki c ile gösterilir ve

$$c = c' + a \cdot b \tag{2.55}$$

formülü ile hesaplanır.

2.14.4 Düzenleyici Etki

Düzenleyici (moderatör) etki, bir bağımsız değişken ile bağımlı değişken arasındaki ilişkiye etki eden düzenleyici adı verilen üçüncü bir değişkenin etkisini ifade eder[12]. Bu etki, bağımsız ve bağımlı değişken arasındaki ilişkinin gücünü veya yönünü, düzenleyici değişkenin düzeyine veya değerine bağlı olarak değiştirdiğinde ortaya çıkar[12]. YEM'de, düzenleyici etkileri, modelde bağımsız değişken ile düzenleyici değişkeni arasındaki etkileşim terimlerinin dahil edilmesi yoluyla incelenebilir[12]. Düzenleyici etkilerin incelenmesi, farklı koşulların veya alt grupların değişkenler arasındaki ilişkileri nasıl etkileyebileceğini anlamak için önemlidir ve bu ilişkilerin hangi sınır koşullarında geçerli olduğuna dair bilgi sağlar[12].



Şekil 2.14 Düzenleyicilik etkisinin istatistiksel diyagramı

3
UYGULAMA

3.1 Uygulama Verisi

Uygulamada, 2016-2017 yılında İstanbul Üniversitesi Diş Hekimliği Fakültesi Pedodonti Anabilim dalında tedavi görmüş, yaşları 3 ve 15 arasında değişen, 118 tek yumurta ikizi, 286 çift yumurta ikizi olmak üzere toplam 404 çocuğa ait veriler üzerinde çalışılmıştır[131]. Veri, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) tarafından 214S284 numaralı hibe ile finanse edilen bir proje için toplanmıştır[131]. Toplanan verilerde, çocuklara ait cinsiyet, biberon ve emzik kullanımı, kötü ağız alışkanlıkları ve bazı ortodontik rahatsızlıkların görülüp görülmemesine dair kategorik değişkenler bulunmaktadır. Tez uygulamasında kullanılan değişkenler aşağıdaki ayrıntılı bir şekilde belirtilmiştir.

1. **Cinsiyet:** Nominal değişkendir. 0= Kız, 1= Erkek şeklinde kodlanmıştır.
2. **Zigosite:** Nominal değişkendir. 0= Tek Yumurta İkizi, 1= Çift yumurta İkizi şeklinde kodlanmıştır.
3. **Biberon Kullanım Süresi:** Ordinal Değişkendir 1= 0-24 ay kullanım, 2 = 24 ay ve üzerinde kullanım olarak kodlanmıştır.
4. **Hamilelik süresi:** Haftalık cinsten nümerik değişkendir.
5. **Doğum Kilosu:** Ordinal değişkendir. 1= 1500 gram altı, 2 = 1500-2500 gram arası, 3=2501-4000 gram arası şeklinde kodlanmıştır.
6. **Süt Dişi Sürme Yaşı:** Aylık cinsten nümerik değişkendir.
7. **Doğum Şekli:** Nominal değişkendir. 0= Normal Doğum, 1= Sezaryen şeklinde kodlanmıştır.
8. **6 ay Anne Sütü Emme:** Nominal değişkendir. 0=Mevcut Değil, 1=Mevcut olarak kodlanmıştır.
9. **Emzik Kullanma:** Nominal değişkendir. 0=Mevcut Değil, 1=Mevcut olarak kodlanmıştır.
10. **Parmak Emme:** Nominal değişkendir. 0=Mevcut Değil, 1=Mevcut olarak kodlanmıştır.

11. **Tırnak Yeme:** Nominal değişkendir. 0=Mevcut Değil, 1=Mevcut olarak kodlanmıştır.
12. **Ağızdan Solunum:** Nominal değişkendir. 0=Mevcut Değil, 1=Mevcut olarak kodlanmıştır.
13. **Atipik Yutkunma:** Nominal değişkendir. 0=Mevcut Değil, 1=Mevcut olarak kodlanmıştır.
14. **Open bite:** Nominal değişkendir. 0=Mevcut Değil, 1=Mevcut olarak kodlanmıştır.

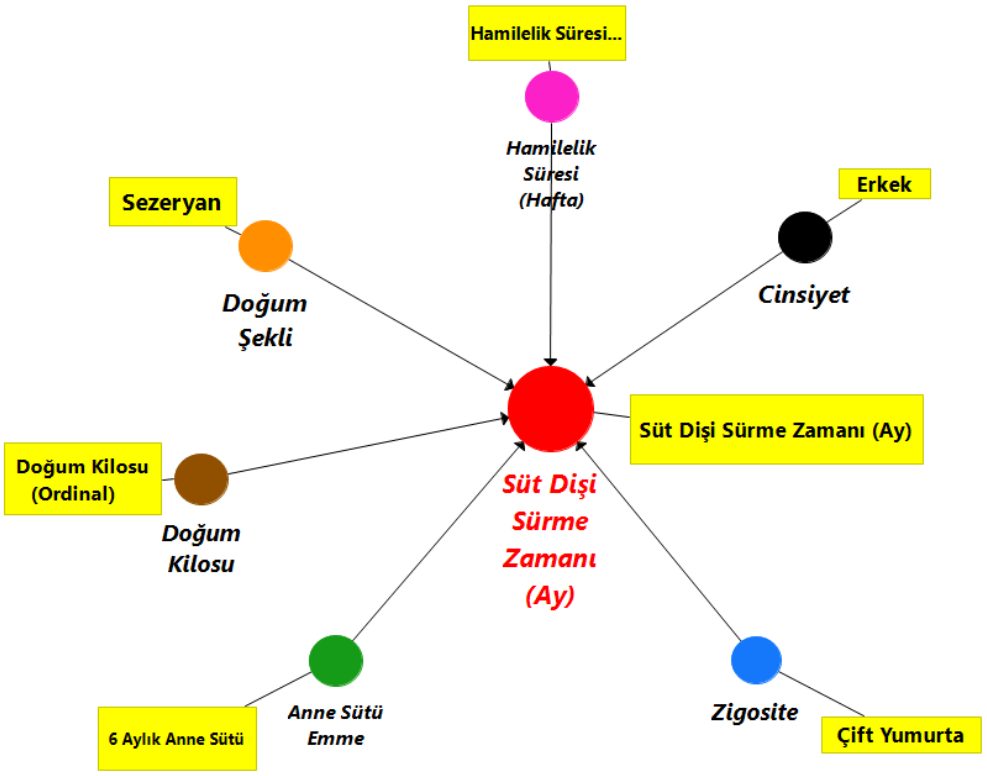
Doğrusal ve lojistik regresyonları için SPSS 25 programı [132], KEKK-YEM modellemeri için ise SmartPLS 3 programı [133], Sobel testi için "Calculation for the Sobel Test: An Interactive Calculation Tool for Mediation Tests" isimli çevrimiçi araç [134], düzenleyici analizi için ise PROCESS 3.5 makrosu [116] kullanılmıştır. Yapısal eşitlik modelleri örneklem büyüklüğüne bağlı olarak hesaplamanın mümkün olduğu durumlarda en az 5000, mümkün olmadığı durumlarda 1000 yeniden örnekleme ile çalışılmıştır.

Uygulamada 3 farklı model şekli çalışılmıştır. Bunlar;

A. İkili Bağımsız Değişken İle Nümerik Bağımlı Değişkenin Bulunduğu Model
B. İkili Bağımsız Değişken, İkili Bağımlı Değişken ve İkili Aracı Değişkenin Bulunduğu Aracı Değişken Modeli
C. İkili Bağımsız Değişken, İkili Bağımlı Değişken ve İkili Düzenleyici Değişkenin Bulunduğu Düzenleyici Değişken Model

3.2 İkili Bağımsız Değişken İle Nümerik Bağımlı Değişkenin Bulunduğu Model İncelemesi

Çalışmanın uygulama kısmında, ilk olarak ikili değişkenlerin nümerik değişkenlerin üzerindeki etkisinin araştırıldığı KEKK-YEM modellemesinin incelemesi yapılmıştır. Bunun için; doğum şekli, hamilelik süresi, cinsiyet, zigosite, 6 ay boyunca anne sütü emme, doğum kilosu değişkenlerinin bağımsız değişken olarak, süt dişi sürme yaşının ise bağımlı değişken olarak kullanıldığı bir model tasarlanmıştır.



Şekil 3.1 Birinci yapısal eşitlik modelininin smartpls gösterimi

Model, 2 farklı örneklem büyüklüğü durumuna göre tekrar tekrar çalıştırılmıştır ve elde edilen bulgular 2 duruma göre ayrı ayrı değerlendirilmiştir. Değerlendirme için kullanılan modeller Model A1, Model A2 şeklinde isimlendirilmiştir.

3.2.1 Örneklem Büyüklüğünün Grup Başına 30'un Altında Kaldığı Küçük Örneklemler İçin Yöntem Karşılaştırılması

İlk olarak ikili nominal değişkenlerde deney ya da kontrol gruplarından bir tanesinin örneklem büyüklüğünün 30'un altında kaldığı, toplam örneklem büyüklüğünün ise 60 olduğu küçük örneklemli bir çalışma modeli simüle edilmiştir. Örneklem seçimleri Excel programı yardımıyla rassal şekilde gerçekleştirilmiştir. Çalışmaya ait örneklemin tanımlayıcı istatistikleri Tablo 3.1'de gösterilmiştir.

Tablo 3.1 Model A1 için tanımlayıcı istatistikler

N=60			
Değişken	**Grup**	n	%
Cinsiyet	**Kız**	35	58,3
	Erkek	25	41,7
Zigosite	**Tek Yumurta**	30	50,0
	Çift Yumurta	30	50,0
Doğum Şekli	**Normal**	8	13,3
	Sezaryen	52	86,7
Anne Sütü (6Ay)	**Hayır**	20	33,3
	Evet	40	66,7

Model öncelikle doğrusal regresyon yöntemi ile test edilmiş daha sonra ise KEKK-YEM modellemesi çalışılmıştır. Elde edilen sonuçlar Tablo 3.2'de gösterilmiştir.

Tablo 3.2 Model A1'dan elde edilen katsayılara ait istatistikler

Yöntem	**Tahminci Değişken**	**β**	**Std. β**	**t**	**p**
Doğrusal Regresyon Modeli	**Model Sabiti**	9,607		3,478	0,001
	Cinsiyet	0,622	0,152	1,069	0,290
	Zigosite	-0,088	-0,022	-0,165	0,870
	Anne Sütü (6 Ay)	-0,104	-0,024	-0,175	0,862
	Hamilelik Süresi	0,001	0,003	0,017	0,987
	Doğum Şekli	-1,565	-0,263	-2,000	0,051
	Doğum Kilosu	-0,567	-0,178	-1,150	0,255
	Tahminci Değişken	**Orijinal Örneklem (β)**	**Örneklem Ort. (β)**	**t**	**p**
KEKK-YEM Modeli	**Doğum Kilosu**	-0,178	-0,162	1,095	0,274
	Zigosite	-0,022	-0,019	0,155	0,877
	Cinsiyet	0,152	0,149	1,055	0,292
	Doğum Şekli	-0,263	-0,255	1,673	0,095
	Anne Sütü (6 Ay)	-0,024	-0,021	0,185	0,854
	Hamilelik Süresi	0,003	-0,002	0,015	0,988

Tablo 3.2 incelendiğinde hem doğrusal regresyon yöntemi ile hem de KEKK-YEM modeli ile elde edilen katsayılarının (doğrusal regresyon için standartlaştırılmış B değerleri, KEKK-YEM için ise orijinal örneklem B değerleri) birbirine eşit olduğu,

bu parametrelere ait p anlamlılık değerlerinin ise birbirlerine çok yakın olduğu görülmektedir. Aynı şekilde orijinal örneklem B değerleri ile yeniden örnekleme ile elde edilen örneklem ortalaması B değerleri arasındaki farklarda göz ardı edilebilecek seviyededir. Diğer bir deyişle yan (bias) miktarı düşüktür.

Bu durumda KEKK-YEM modellemesi yol analizi ile doğrusal regresyon analizinin küçük örneklemlerde benzer sonuçlar verebildiğini söylemek mümkündür.

Bununla birlikte doğum şekli değişkeni için elde edilen p değerleri kontrol edildiğinde, KEKK-YEM'in hesapladığı olasılık değerinin doğrusal regresyon modelinden elde edilen p değerinin yaklaşık 2 katı olduğu görülmektedir. Bu durum, model daha geniş bir örneklem ile çalıştırılarak sonuçlar tekrar kontrol edilmiştir.

3.2.2 Örneklem Büyüklüğünün Grup Başına 30'un Üstünde Olduğu Durumlar İçin Yöntem Karşılaştırılması

Şekil 3.1'de gösterilen KEKK-YEM modeli daha geniş ve kategorik değişkenler için grup başına düşen örneklem büyüklüğünün 30'un üzerinde olduğu bir örneklem ile tekrar çalıştırılmıştır. Örnekleme ait frekans bilgileri Tablo 3.3'te gösterilmiştir.

Tablo 3.3 Model A2 için tanımlayıcı istatistikler

N=404			
Değişken	**Grup**	**n**	**%**
Cinsiyet	**Kız**	200	49,5
	Erkek	204	50,5
Zigosite	**Tek Yumurta**	118	29,2
	Çift Yumurta	286	70,8
Doğum Şekli	**Normal**	47	11,6
	Sezaryen	357	88,4
Anne Sütü (6 Ay)	**Hayır**	133	32,9
	Evet	251	62,1

Model daha önce olduğu gibi öncelikle doğrusal regresyon yöntemi ile test edilmiş daha sonra ise KEKK-YEM modellemesi çalışılmıştır. Elde edilen sonuçlar Tablo 3.4'te gösterilmiştir.

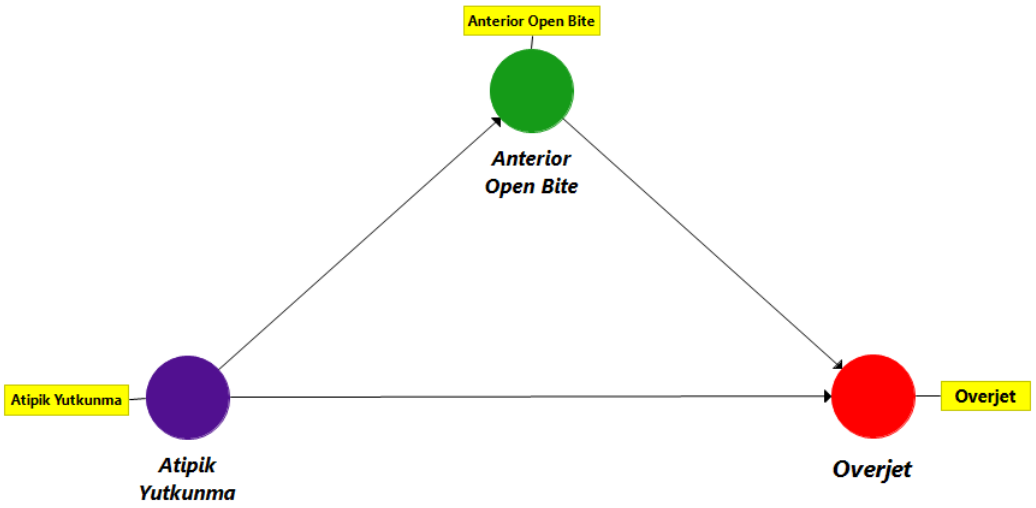
Tablo 3.4 Model A2'den elde edilen katsayılara ait istatistikler

Yöntem	Tahminci Değişken	β	Standart β	t	p
Doğrusal Regresyon Modeli	Model Sabiti	8,056		5,451	0,000*
	Cinsiyet	0,346	0,072	1,431	0,153
	Zigosite	-0,582	-0,112	-2,230	0,026*
	Anne Sütü (6 Ay)	0,298	0,059	1,174	0,241
	Hamilelik Süresi (Hafta)	0,040	0,053	0,901	0,368
	Doğum Şekli	-0,682	-0,091	-1,803	0,072
	Doğum Kilosu	-0,809	-0,215	-3,620	0,000*
	Tahminci Değişken	Örneklem Ort. (Std. β)	Orijinal Örneklem (Std. β)	t	p
KEKK-YEM Modeli	Doğum Kilosu	-0,210	-0,214	2,964	0,003*
	Zigosite	-0,115	-0,118	2,23	0,026*
	Cinsiyet	0,068	0,071	1,506	0,132
	Doğum Şekli	-0,086	-0,086	1,411	0,158
	Anne Sütü (6 Ay)	0,057	0,059	1,229	0,219
	Hamilelik Süresi (Hafta)	0,0610	0,063	0,967	0,333

Tablo 3.4 incelendiğinde, doğrusal regresyon ve KEKK-YEM modellemesi ile elde edilen katsayılarının ve p anlamlılık değerlerinin, daha önceki gibi birbirlerine çok yakın olduğu görülmektedir. Doğum şekli değişkeni için elde edilen sonuçlar kontrol edildiğinde ise her iki yöntem için de B katsayılarının sıfıra yaklaştığı, p değerlerinin ise yükseldiği görülmektedir. Doğrusal regresyon ile elde edilen p değeri 0,051'den 0,072'ye, KEKK-YEM modellemesi ile elde edilen p değeri ise 0,095'ten 0,158'e yükselmiştir. Bu durumda, değerler arasındaki farkın örneklem büyüklüğünün 30'un altında ya da üstünde olması ile ilişkili olmadığı söylenebilir. Diğer taraftan Tablo 3.1 ve Tablo 3.3 incelendiğinde sezaryen grubunun örneklem büyüklüğünün normal doğum yapanlara göre yaklaşık 8 kat daha fazladır ve gruplar arasındaki bu fark diğer değişkenlere göre çok daha yüksektir. Grupların örneklem büyükleri arasında bu derece asimetrik bir durum olduğunda, KEKK-YEM'in klasik doğrusal regresyona göre α tipi hataya (1. Tip hataya) daha dirençli olması beklenebilir. Diğer bir ifade ile böyle bir durumda KEKK-YEM'in istatistiki güven düzeyi daha yüksek olabilir.

3.3 İkili Bağımsız Değişken, İkili Bağımlı Değişken ve İkili Aracı Değişkenin Bulunduğu Aracılık Modeli İncelemesi

Uygulamanın ikinci bölümünde, bir ikili değişkenin aynı anda hem bağımlı hem bağımsız değişken olarak kullanıldığı aracılık modellemeleri için mevcut olan yöntemler karşılaştırılmıştır. Bunun için bağımsız değişken olarak atipik yutkunma, aracı değişken olarak open bite rahatsızlığı, bağımlı değişken olarak ise overjet rahatsızlığı kullanılmıştır. Tasarlanan model Şekil 3.2'de gösterilmiştir.



Şekil 3.2 Kurulan aracılık modelinin smartpls gösterimi

Model, bir önceki uygulamadaki gibi 3 farklı örneklem büyüklüğü durumuna göre tekrar tekrar çalıştırılmıştır ve elde edilen bulgular 3 duruma göre ayrı ayrı değerlendirilmiştir. Yapılan değerlendirmeler; sırasıyla Model B1, Model B2 ve Model 3 şeklinde isimlendirilmiştir.

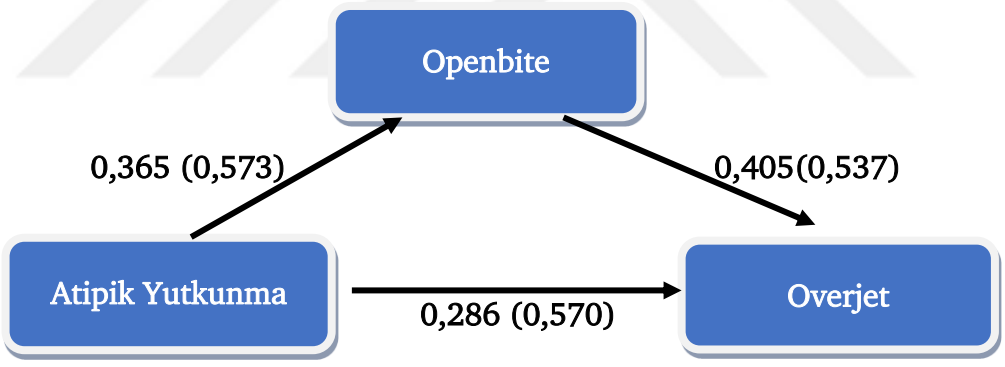
3.3.1 Örneklem Büyüklüğünün Grup Başına 30'un Altında Kaldığı Küçük Örneklemler İçin Yöntem Karşılaştırılması

İlk olarak ikili nominal değişkenlerde deney ya da kontrol gruplarından bir tanesinin örneklem büyüklüğünün 30'un altında kaldığı, toplam örneklem büyüklüğünün ise 60 olduğu küçük örneklemli bir çalışma modeli simüle edilmiştir. Çalışmaya ait örneklemin tanımlayıcı istatistikleri Tablo 3.5'te gösterilmiştir.

Tablo 3.5 Model B1 için tanımlayıcı istatistikler

N=60			
Değişken	**Grup**	**n**	**%**
Atipik Yutkunma	**Hayır**	42	70,0
	Evet	18	30,0
Openbite	**Hayır**	37	61,7
	Evet	23	38,3
Overjet	**Hayır**	30	50,0
	Evet	30	50,0

Öncelikle lojistik regresyon yöntemi ile 2 ayrı model kurulmuştur. İlk modelde bağımsız değişken olarak kullanılan atipik yutkunmanın aracı değişken olarak kullanılan openbite rahatsızlığı üzerindeki etkisi kontrol edilmiştir. İkinci regresyon modelinde ise bu iki değişkenin aynı anda her ikisi de tahminci değişken, overjet rahatsızlığı ise bağımlı değişken olarak kullanılmıştır. Elde edilen B katsayıları ve o katsayılara ait standart hatalar (parantezlerin içerisinde) Şekil 3.3'te gösterilmiştir.

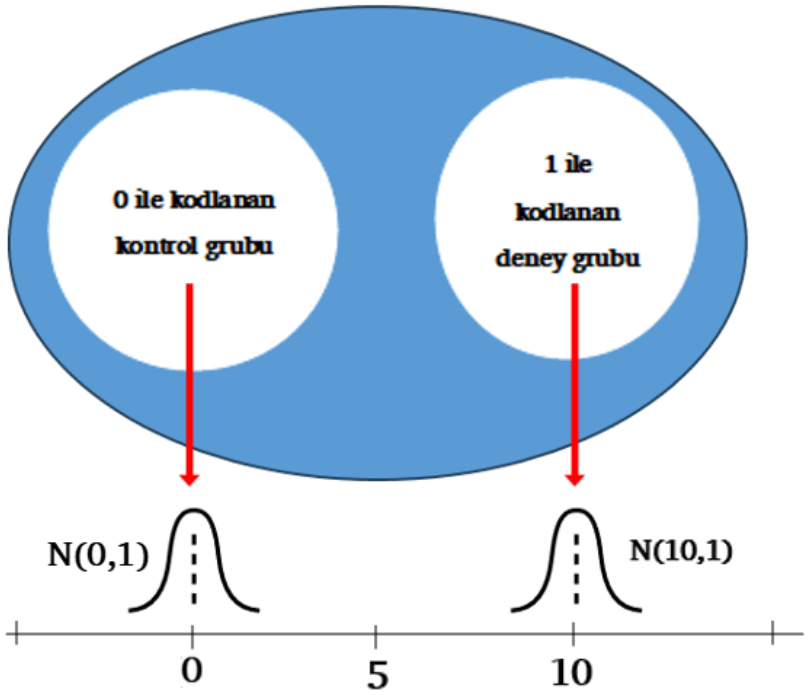


Şekil 3.3 Model B1 için lojistik regresyon ile elde edilen b katsayılar ve standart hataları

Model aynı zamanda KEKK-YEM modellemesi ile test edilmiştir. KEKK-YEM modellemesinde bağımlı değişkenin kategorik bir değişken olması, yöntem metodu olarak uygun olmadığı için ve kurulan modeldeki bağımlı değişken (openbite ve overjet) bir ikili değişken olduğu için modeldeki 3 kategorik değişken de hem dummy kodlama ile hem de nümerik kodlama ile değerlendirilmiştir.

Nümerik kodlama için veriye şu şekilde bir dönüşüm uygulanmıştır. Dummy kodlamaya göre 0 ile kodlanan kontrol grup için ortalaması 0, standart sapması

olan 1 olan bir standart normal dağılım verisi türetilmiştir. Dummy kodlamaya göre 1 ile kodlanan vakaların bulunduğu deney grubu için ise standart sapması 1, ortalaması ise sırasıyla 2,3,4,….,30 olan çok sayıda normal dağılım verisi türetilerek farklı denemeler yapılmış ve elde edilen sonuçların m=10 noktasından sonra stabil hale geldiği görülmüştür ve dönüşümün bu şekilde yapılmasında karar kılınmıştır.



Şekil 3.4 İkili değişkenlere uygulanan dönüştürme

KEKK-YEM modeli, kullanılan 3 değişkenin tamamı hem kategorik değişkene hem de nümerik veriye göre ayrı ayrı kodlanarak tekrar çalıştırılmıştır. 3 farklı yöntemden elde edilen sonuçlar Tablo 3.6'da gösterilmiştir.

Tablo 3.6 Model B1'den elde edilen katsayılara ait istatistikler

Yöntem	Tahmin Edilen Değişken	Tahminci Değişken	β	OR	Wald	p
İkili Lojistik Regresyon (1. Model)	Openbite	Sabit	-0,588	0,556	3,332	0,068
		Atipik Yutkunma	0,365	1,440	0,405	0,525
İkili Lojistik Regresyon (2. Model)	Overjet	Sabit	-0,240	,786	0,433	0,511
		Openbite	0,405	1,499	0,567	0,451
		Atipik Yutkunma	0,286	1,331	0,252	0,616
Yöntem	**Tahmin Edilen Değişken**	**Tahminci Değişken**	**Örneklem Ort. (Std. β)**	**Orijinal Örneklem (Std. β)**	**t**	**p**
KEKK-YEM Modeli (Dummy Kodlama İle)	Openbite	Atipik Yutkunma	0,075	0,082	0,648	0,517
	Overjet	Openbite	0,100	0,098	0,746	0,456
	Overjet	Atipik Yutkunma	0,066	0,065	0,511	0,610
KEKK-YEM Modeli (Nümerik Kodlama İle)	Openbite	Atipik Yutkunma	0,105	0,101	0,768	0,443
	Overjet	Openbite	0,088	0,086	0,649	0,517
	Overjet	Atipik Yutkunma	0,081	0,082	0,606	0,545

Tablo 3.6 kontrol edildiğinde lojistik regresyon sonuçlarına göre, kurulan her iki modelde de bağımsız, aracı ve bağımlı değişkenler (atipik yutkunma, openbite ve overjet) arasında istatistiki olarak anlamlı bir ilişki görülmemektedir($p>0,05$). Buna göre modelde bir aracılık rolünden bahsedilmez. KEKK-YEM modellemesi denemelerinde de ikili lojistik regresyon modeline paralel şekilde değişkenler arasındaki hiçbir ilişki için istatistiksel anlamlılık yakalanmamıştır ($p>0,05$).

Sonuçlar incelendiğinde, değişkenler dummy kodlama ile kodlandığında KEKK-YEM ikili lojistik regresyon ile aynı anlamlılık değerlerini verdiği görülmektedir. Diğer yandan nümerik kodlama yapıldığında ise atipik yutkunma → openbite ve atipik yutkunma → overjet için yaklaşık 0,06 kadar daha düşük, openbite→overjet

yolu için ise 0,06 kadar daha yüksek değerler hesaplandığı görülmektedir. Aracılık analizi için ise yapılan Sobel, Aroian, Goodman ve KEKK-YEM testlerinin sonuçları ise Tablo 3.7'de gösterilmiştir.

Tablo 3.7 Model B1 için aracılık analizi sonuçları

Atipik Yutkunma → Openbite → Overjet		
Yöntem	Z	p
Sobel	0,487	0,627
Aroian	0,341	0,732
Goodman	-	-
Yöntem	**t**	**p**
KEKK-YEM (Dummy Kodlama İle)	0,345	0,730
KEKK-YEM (Nümerik Kodlama İle)	0,344	0,731

Yapılan aracılık analizleri sonucunda Aroin testi ile ve KEKK-YEM ile hem dummy hem nümerik kodlama yapıldığında aynı anlamlılık değerinin (0,73) elde edildiği görülmüştür. Sobel testi sonucundaysa diğer yöntemlere göre yaklaşık 0,10 kadar daha düşük bir p değeri (0,63) hesaplamıştır. Goodman testinden ise sonuç elde edilememiştir.

Model, örneklem büyüklüğünün grup başına 30'un altında kaldığı, diğer bir ifade ile gruplar arasında frekans farkının yüksek seviyede olduğu büyük örneklemli değişkenler ile tekrar çalıştırılmıştır.

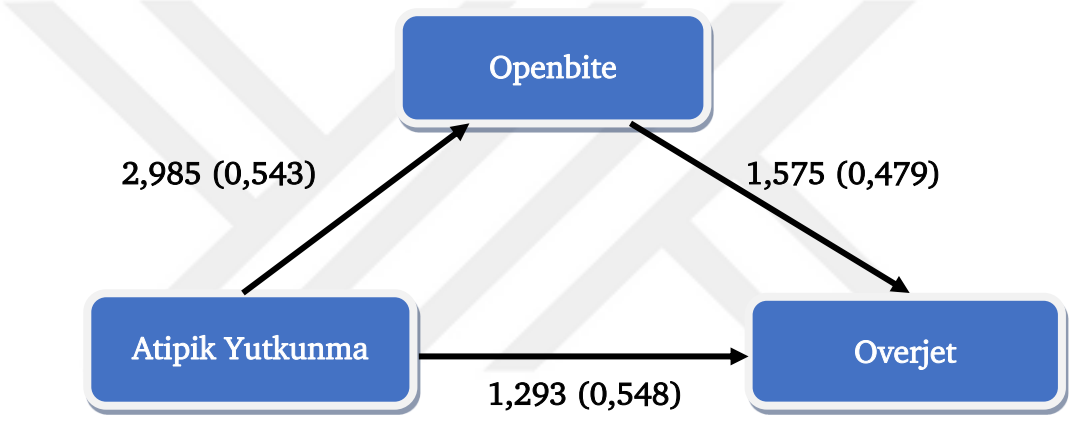
3.3.2 Örneklem Büyüklüğünün Grup Başına 30'un Altında Kaldığı Büyük Örneklem İçin Yöntem Karşılaştırılması

Uygulamanın bu kısmında, 2. aracılık modellemesi için mevcut olan 404 gözlemin tamamı kullanılmıştır. Modelde kullanılan 3 değişkene ait frekans istatistikleri Tablo 3.8'de verilmiştir.

Tablo 3.8 Model B2 için tanımlayıcı istatistikler

N=404			
Değişken	**Grup**	**n**	**%**
Atipik Yutkunma	**Hayır**	386	95,5
	Evet	18	4,5
Openbite	**Hayır**	381	94,3
	Evet	23	5,7
Overjet	**Hayır**	331	81,9
	Evet	73	18,1

Bir önceki uygulamada da yapıldığı üzere, ilk olarak lojistik regresyon yöntemi ile 2 ayrı model kurulmuştur. Elde edilen B katsayıları ve o katsayılara ait standart hatalar Şekil 3.5'te gösterilmiştir.



Şekil 3.5 Model B2 için lojistik regresyon ile elde edilen b katsayılar ve standart hataları

Model KEKK-YEM modellemesi ile, hem dummy kodlama ile hem de nümerik kodlama uygulanarak (Şekil 3.4) değerlendirilmiştir. 3 farklı yöntemden elde edilen sonuçlar Tablo 3.9'da gösterilmiştir.

Tablo 3.9 Model B2'den elde edilen katsayılara ait istatistikler

Yöntem	Tahmin Edilen Değişken	Tahminci Değişken	β	OR	Wald	p
Lojistik Regresyon Modeli	Openbite	Sabit	-3,208	0,040	148,384	0,000*
		Atipik Yutkunma	2,985	19,787	30,270	0,000*
	Overjet	Sabit	-1,726	0,178	143,934	0,000*
		Openbite	1,575	4,830	10,830	0,001*
		Atipik Yutkunma	1,293	3,645	5,562	0,018*
Yöntem	**Tahmin Edilen Değişken**	**Tahminci Değişken**	**Örneklem Ort. (Std. β)**	**Orijinal Örneklem (Std. β)**	**t**	**p**
KEKK-YEM Modeli (Dummy Kodlama İle)	Openbite	Atipik Yutkunma	0,356	0,361	3,528	0,000*
	Overjet	Openbite	0,194	0,195	2,648	0,008*
	Overjet	Atipik Yutkunma	0,140	0,140	1,910	0,056
KEKK-YEM Modeli (Nümerik Kodlama İle)	Openbite	Atipik Yutkunma	0,291	0,292	3,411	0,001*
	Overjet	Openbite	0,196	0,195	3,117	0,002*
	Overjet	Atipik Yutkunma	0,14	0,14	2,14	0,032*

Tablo 3.9 kontrol edildiğinde lojistik regresyon sonuçlarına göre, kurulan her iki modelde de bağımsız, aracı ve bağımlı değişkenler (atipik yutkunma, openbite ve overjet) arasında istatistiki olarak anlamlı bir ilişki görülmektedir($p<0,05$). Buna göre modelde aracı değişken olarak kullanılan openbite değişkeni kısmi aracılık rolüne sahip olabilir ve atipik yutkunma ile overjet arasındaki dolaylı etki kontrol edilmelidir.

KEKK-YEM modellemesi, lojistik regresyon ile karşılaştırıldığında atipik yutkunma →openbite ve openbite→overjet yolları için aynı sonuçlar edilirken, atipik yutkunma→overjet yolu için farklı sonuçlar elde edilmiştir.

İkili değişkenler 0,1 şeklinde dummy kodlama ile kodlandığında, KEKK-YEM modelinde atipik yutkunma→overjet yolu için elde edilen olasılık değeri, ikili lojistik regresyon modelinde elde edilen değere ($p=0,018$) oranla yaklaşık 3

katına çıkmıştır ve 0,05 anlamlılık düzeyinin de üzerine yükselmiştir (p=0,056). Bu durumda modelde istatistiki olarak anlamlı bir dolaylı etki var ise aracı değişken olarak kullanılan openbite değişkeni tam aracılık rolüne sahip olabilir.

Modeldeki ikili değişkenler nümerik olarak kodlandığında ise KEKK-YEM modelinde atipik yutkunma→overjet yolu için elde edilen olasılık değeri, ikili lojistik regresyon modelinde elde edilen değere oranla yaklaşık 2 katına çıkmıştır ve fakat 0,05 anlamlılık düzeyinin altında kalmıştır (p=0,036). Bu durumda modelde istatistiki olarak anlamlı bir dolaylı etki var ise aracı değişken olarak kullanılan openbite değişkeni tam aracılık rolüne sahip olabilir.

Elde edilen bu sonuçlara dayanarak grup başına örneklem büyüklüğünün 30'un altında kaldığı ve gruplar arasında frekans farkının yüksek olduğu durumlarda, ikili değişkenleri dummy yerine nümerik kodlama ile kodlamanın, tip 2 hatadan (β hatasından) kaçınmaya diğer bir ifadeyle testin gücünü yükseltmeye yardımcı olabilir.

Aracılık analizi için ise yapılan Sobel, Aroian, Goodman ve KEKK-YEM testlerinin sonuçları ise Tablo 3.10'da gösterilmiştir.

Tablo 3.10 Model B2 için aracılık analizi sonuçları

Atipik Yutkunma → Openbite → Overjet		
Test	**Z**	**p**
Sobel	2,822	0,004*
Aroian	2,788	0,005*
Goodman	2,856	0,004*
Test	**t**	**p**
KEKK-YEM (Dummy Kodlama İle)	2,123	0,034*
KEKK-YEM (Nümerik Kodlama İle)	2,235	0,025*

Sobel aracılık analizleri sonucunda, üç testin tamamına göre atipik yutkunma ve overjet arasındaki dolaylı etki %1 anlamlılık düzeyinde istatistiksel olarak anlamlıdır(p<0,01). Buna göre, atipik yutkunma ve overjet arasındaki ilişki istatistiki olarak anlamlı olduğu için, modelde openbite değişkeni "kısmi aracılık rolüne" sahiptir.

KEKK-YEM sonuçlarına göre ise, değişkenler hem dummy hem nümerik kodlama ile kodlandığında atipik yutkunma ve overjet arasındaki dolaylı etki %5 anlamlılık düzeyinde istatistiksel olarak anlamlıdır($p<0,05$). Ancak openbite değişkeninin aracı değişken olarak oynadığı rol, dummy kodlama yapıldığında "tam aracılık", nümerik kodlama yapıldığında "kısmi aracılık" olarak yorumlanmak durumundadır. Diğer taraftan dolaylı etki için Sobel ve KEKK-YEM sonuçları arasındaki farklılaşmanın, modeldeki bağımlı ve bağımsız değişkenler arasındaki ilişkinin ikili lojistik regresyon ve KEKK-YEM arasında farklı sonuçlar vermiş olması ile ilgilidir.

Model üçüncü ve son kez, grup başına örneklem büyüklüğünün grup başına 30'un üstünde olduğu fakat frekans farkının yine yüksek seviyede olduğu daha büyük bir örneklem ile tekrar çalıştırılmıştır.

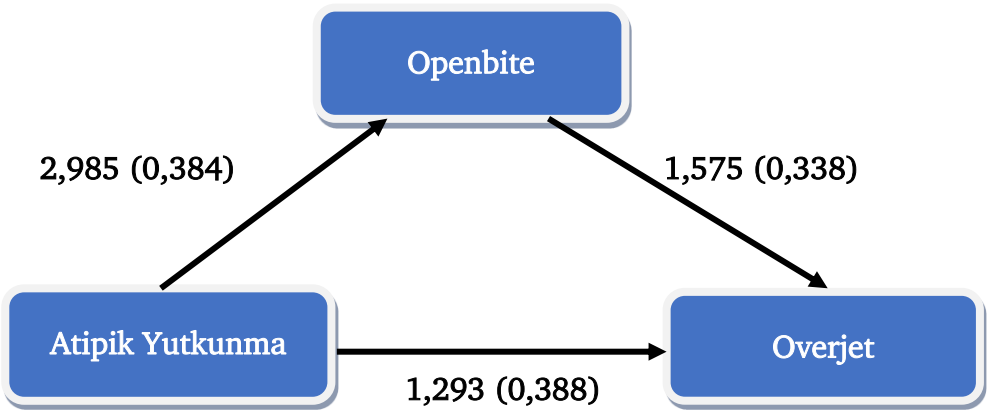
3.3.3 Örneklem Büyüklüğünün Grup Başına 30'un Üstünde Olduğu Büyük Örneklemler İçin Yöntem Karşılaştırılması

Bu simülasyon için, mevcut 404 gözlem kopyalanarak örneklem büyüklüğü genişletilmiştir ve grup başına düşen en küçük örneklem büyüklüğü 30'un üstüne çıkarılmıştır. Frekans istatistikleri Tablo 3.11'de verilmiştir.

Tablo 3.11 Model B3 için tanımlayıcı istatistikler

N=808			
Değişken	**Grup**	**n**	**%**
Atipik Yutkunma	**Hayır**	772	95,5
	Evet	36	4,5
Openbite	**Hayır**	762	94,3
	Evet	46	5,7
Overjet	**Hayır**	662	81,9
	Evet	146	18,1

Bir önceki uygulamada da yapıldığı üzere, ilk olarak lojistik regresyon yöntemi ile 2 ayrı model kurulmuştur. Elde edilen B katsayıları ve o katsayılara ait standart hatalar Şekil 3.6'da gösterilmiştir.



Şekil 3.6 Model B3 için lojistik regresyon ile elde edilen b katsayılar ve standart hataları

Model KEKK-YEM modellemesi ile, hem dummy kodlama ile hem de nümerik kodlama uygulanarak (Şekil 3.4) değerlendirilmiştir. 3 farklı yöntemden elde edilen sonuçlar Tablo 3.12'de gösterilmiştir.

Tablo 3.12 Model B3'ten elde edilen katsayılara ait istatistikler

Yöntem	Tahmin Edilen Değişken	Tahminci Değişken	β	OR	Wald	p
Lojistik Regresyon Modeli	Openbite	Sabit	-3,208	0,040	296,768	0,000*
		Atipik Yutkunma	2,985	19,787	60,540	0,000*
	Overjet	Sabit	-1,726	,178	287,868	0,000*
		Openbite	1,293	3,645	11,124	0,001*
		Atipik Yutkunma	1,575	4,830	21,661	0,000*
Yöntem	**Tahmin Edilen Değişken**	**Tahminci Değişken**	**Örneklem Ort. (Std. β)**	**Orijinal Örneklem (Std. β)**	**t**	**p**
KEKK-YEM Modeli (Dummy Kodlama İle)	Openbite	Atipik Yutkunma	0,359	0,361	5,116	0,000
	Overjet	Openbite	0,194	0,195	3,933	0,000
	Overjet	Atipik Yutkunma	0,140	0,140	2,710	0,007
KEKK-YEM Modeli (Nümerik Kodlama İle)	Openbite	Atipik Yutkunma	0,291	0,292	4,848	0,000
	Overjet	Openbite	0,197	0,195	4,347	0,000
	Overjet	Atipik Yutkunma	0,139	0,140	2,984	0,003

İkili lojistik regresyon sonuçlarına göre, kurulan her iki modelde de bağımsız, aracı ve bağımlı değişkenler (atipik yutkunma, openbite ve overjet) arasında istatistiki olarak yüksek derece anlamlı bir ilişki görülmektedir($p<0,01$). Buna göre modelde aracı değişken olarak kullanılan openbite değişkeni kısmi aracılık rolüne sahip olabilir ve atipik yutkunma ile overjet arasındaki dolaylı etki kontrol edilmelidir.

Diğer taraftan hem dummy hem nümerik kodlama şekli ile yapılan KEKK-YEM analizinde, ikili lojistik regresyon ile paralel sonuçlar elde edilmiştir. Modelde test edilen tüm yollar için istatistiki olarak yüksek derece anlamlı ilişki tespit edilmiştir ($p<0,01$).

Tablo 3.13 Model B3 için aracılık analizi sonuçları

Atipik Yutkunma → Openbite → Overjet		
Test	**Z**	**p**
Sobel	3,996	0,000
Aroian	3,972	0,000
Goodman	4,021	0,000
Test	**t**	**p**
KEKK-YEM (Dummy Kodlama İle)	3,135	0,002
KEKK-YEM (Nümerik Kodlama İle)	3,223	0,001

Sobel aracılık analizleri sonucunda, 404 gözlem ile yapılan simülasyon ile hemen hemen aynı sonuçlar elde edilmiştir. Üç testin tamamına göre atipik yutkunma ve overjet arasındaki dolaylı etkiye ait anlamlılık değeri 0,000'dır ve yine atipik yutkunma ve overjet arasındaki ilişki istatistiki olarak anlamlı olduğu için, openbite değişkeni "kısmi aracılık rolüne" sahiptir.

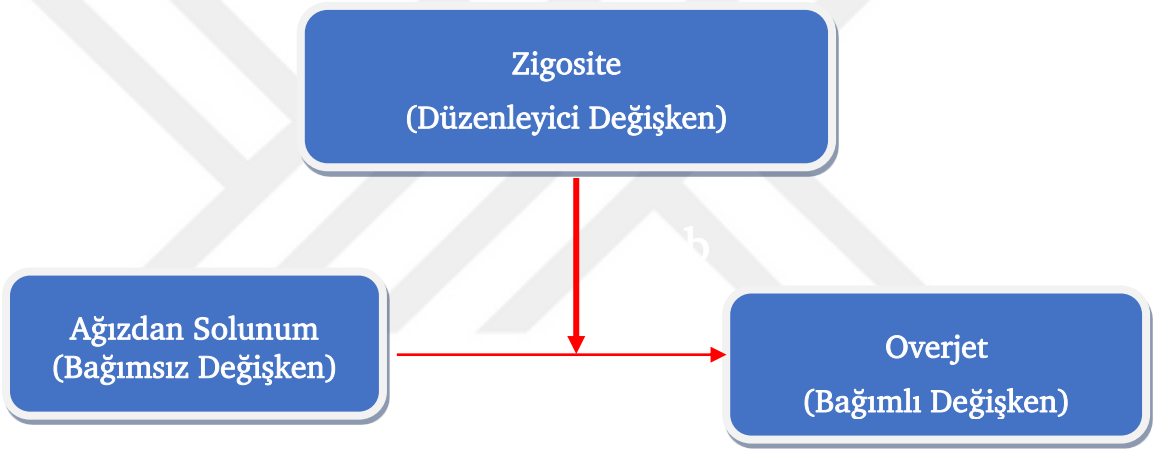
KEKK-YEM sonuçlarına göre ise, değişkenler hem dummy hem nümerik kodlama ile kodlandığında atipik yutkunma ve overjet arasındaki dolaylı etki %1 anlamlılık düzeyinde istatistiksel olarak anlamlıdır($p<0,01$). Her iki kodlama şekli sonucunda da openbite değişkeninin kısmi aracı rolü oynadığı söylenebilir.

Bu bilgilere dayanarak; grup başına düşen örneklem büyüklüklerini 30'dan yüksek olduğu durumlarda, gruplar arası frekans farkı çok yüksek (çarpıklık durumu)

olmasına rağmen, KEKK-YEM'in hem dummy hem de nümerik kodlama tercih edildiğinde sobel testler ile aynı sonuçlar vererek modeldeki dolaylı etkiyi yakalayabildiği beklenen bir durumdur denilebilir.

3.4 İkili Bağımsız Değişken, İkili Bağımlı Değişken ve İkili Düzenleyici Değişkenin Bulunduğu KEKK-YEM Modelinin İncelemesi

Uygulamanın üçüncü bölümünde, 1 adet ikili değişkenin, 2 adet ikili değişken arasındaki ilişkiyi düzenlediği YEM modellemeleri için mevcut olan yöntemler karşılaştırılmıştır. Bunun için bağımsız değişken olarak ağızdan soluma, düzenleyici değişken olarak zigosite, bağımlı değişken olarak ise overjet rahatsızlığı kullanılmıştır. Tasarlanan YEM modeli Şekil 3.7'de gösterilmiştir.



Şekil 3.7 Model C için kavramsal diyagram

Model, bir önceki uygulamadaki gibi 3 farklı örneklem büyüklüğü durumuna göre tekrar tekrar çalıştırılmıştır ve elde edilen bulgular 3 duruma göre ayrı ayrı değerlendirilmiştir. Yapılan değerlendirmeler; sırasıyla Model C1, Model C2 ve Model C3 şeklinde isimlendirilmiştir.

3.4.1 Örneklem Büyüklüğünün Grup Başına 30'un Altında Kaldığı Küçük Örneklemler İçin Yöntem Karşılaştırılması

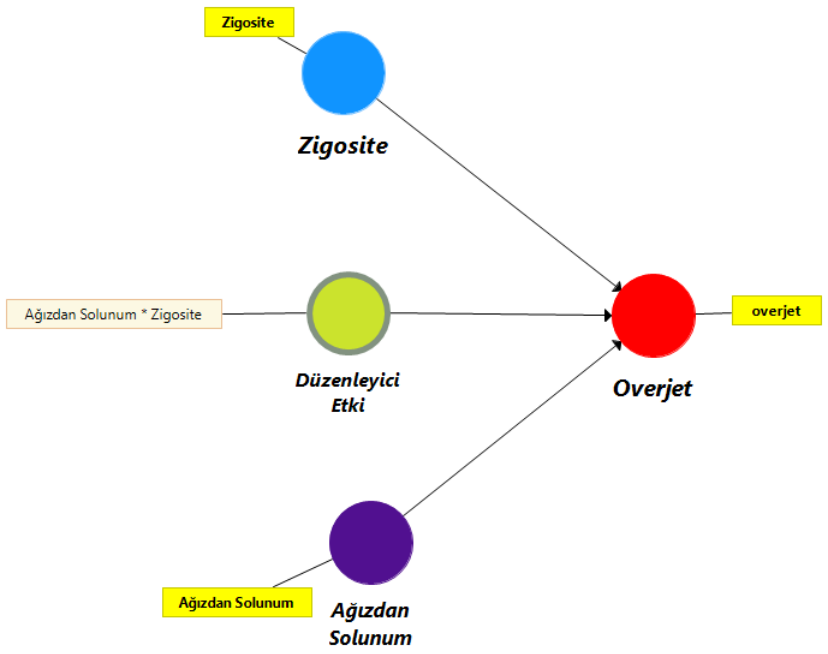
İlk olarak ikili nominal değişkenlerde deney ya da kontrol gruplarından bir tanesinin örneklem büyüklüğünün 30'un altında kaldığı, toplam örneklem büyüklüğünün ise 50 olduğu küçük örneklemli bir çalışma modeli simüle

edilmiştir. Çalışmaya ait örneklemin tanımlayıcı istatistikleri Tablo 3.14'te gösterilmiştir.

Tablo 3.14 Model C1 için tanımlayıcı istatistikler

N=50				
Zigosite	**Değişken**	**Grup**	**n**	**%**
Tek Yumurta İkizi (n=25)	**Ağızdan Soluma**	**Hayır**	12	48,0
		Evet	13	52,0
	Overjet	**Hayır**	16	64,0
		Evet	9	36,0
Çift Yumurta İkizi (n=25)	**Ağızdan Soluma**	**Hayır**	10	40,0
		Evet	15	60,0
	Overjet	**Hayır**	10	40,0
		Evet	15	60,0

Model Hayes'in yazdığı bir makro olan PROCESS ile, dummy kodlaması kullanılarak KEKK-ÇGA ile ve nümerik kodlama kullanılarak KEKK-YEM yöntemleri ile ayrı ayrı analiz edilmiştir. Nümerik kodlama ile kodlanan değişkenlerle kurulan YEM modelinin SmartPLS görünümü Şekil 3.8'de görülebilir.



Şekil 3.8 Model C için tasarlanan KEKK-YEM'in smartpls gösterimi

3 ayrı yöntemden elde edilen sonuçlar Tablo 3.15'te gösterilmiştir.

Tablo 3.15 Model C1'den elde edilen katsayılara ait istatistikler

Yöntem	Düzenleyici Grubu	Tahmin Edilen Değişken	Tahminci Değişken	Koşullu Etki	Std. Hata	z	p
Process Macro	Tek Yumurta İkizi	Overjet	Ağızdan Solunum	-0,474	0,839	-0,565	0,572
	Çift Yumurta İkizi	Overjet	Ağızdan Solunum	-1,520	0,945	-1,608	0,108
Yöntem	Düzenleyici Grubu	Tahmin Edilen Değişken	Tahminci Değişken	Orijinal Örneklem (Std. β)	Std. Sapma	t	p
KEKK-ÇGA (Dummy Kodlama İle)	Tek Yumurta İkizi	Overjet	Ağızdan Solunum	-0,113	0,198	0,572	0,568
	Çift Yumurta İkizi	Overjet	Ağızdan Solunum	-0,333	0,182	1,828	0,068
KEKK-YEM (Nümerik Kodlama İle)		Overjet	Ağızdan Solunum	-0,26	0,135	1,93	0,054
		Overjet	Zigosite	0,29	0,132	2,200	0,028
		Overjet	Düzenleyici Etki	-0,059	0,132	0,449	0,654

Process Macro ve KEKK-ÇGA yöntemleri ile yapılan analizler sonucunda, tek yumurta ikizlerinde ağızdan solunum ve overjet arasındaki negatif bir ilişki için aynı anlamlılık değeri elde edilmiştir(p=0,56). Diğer taraftan çift yumurta ikizleri için Process Macro 0,11 oranında bir p değeri hesaplarken KEKK-ÇGA 0,07'lik bir p değeri hesaplamıştır. Zigosite gruplarından elde B katsayılarının arasındaki farkın anlamlılığın test edildiği düzenleyici değişken analizlerinin sonuçları ise Tablo 3.16'da gösterilmiştir.

Tablo 3.16 Model C1 için düzenleyici değişken analizi sonuçları

Test	Z	p
Process Macro	0,700	0,403
	Katsayı Farkı	p
KEKK-ÇGA (Dummy Kodlama İle)	0,220	0,420
	t	p
KEKK-YEM (Nümerik Kodlama İle)	0,449	0,654

Buna göre Process Macro ve KEKK-ÇGA yöntemlerinin, gruplar arası frekans farkının az olduğu küçük örneklemlerde birbirlerine çok yakın sonuçlar vermesi beklenebilir.

Diğer taraftan değişkenleri birer nominal değişken olarak görmeden nümerik kodlama yapıldığında ve düzenleyicilik etkisi modele eklenen bir düzenleyici değişken ile test edildiğinde (Şekil 3.7) KEKK-YEM'in hesapladığı anlamlılık değeri 0,654 şeklinde diğer yöntemlere kıyasla çok daha yüksek hesaplanmıştır.

Aynı model tasarımı örneklem büyüklüğünün grup başına 30'un altında kaldığı ve gruplar arasında frekans farkının yüksek olduğu daha büyük bir örneklem ile tekrar çalıştırılmıştır.

3.4.2 Örneklem Büyüklüğünün Grup Başına 30'un Altında Kaldığı Büyük Örneklem İçin Yöntem Karşılaştırılması

İkili nominal değişkenlerde deney ya da kontrol gruplarından bir tanesinin örneklem büyüklüğünün 30'un altında kaldığı, toplam örneklem büyüklüğünün ise 245 olduğu büyük örneklemli bir çalışma modeli simüle edilmiştir. Bu çalışma için tek yumurta ikizlerinden 102 adet, çift yumurta ikizlerinden ise 143 adet gözlem rassal bir şekilde seçilmiştir. Çalışmaya ait örneklemin tanımlayıcı istatistikleri Tablo 3.17'de gösterilmiştir.

Tablo 3.17 Model C2 için tanımlayıcı istatistikler

N=245				
Zigosite	**Değişken**	**Grup**	**n**	**%**
Tek Yumurta İkizi (n=102)	**Ağızdan Soluma**	**Hayır**	76	74,5
		Evet	26	25,5
	Overjet	**Hayır**	83	81,4
		Evet	19	18,6
Çift Yumurta İkizi (n=143)	**Ağızdan Soluma**	**Hayır**	121	84,6
		Evet	22	15,4
	Overjet	**Hayır**	123	86,0
		Evet	20	14,0

3 ayrı yöntemden elde edilen sonuçlar Tablo 3.18'de gösterilmiştir.

Tablo 3.18 Model C2'den elde edilen katsayılara ait istatistikler

Yöntem	Düzenleyici Grubu	Tahmin Edilen Değişken	Tahminci Değişken	Koşullu Etki	Std. Hata	z	p
Process Macro	Tek Yumurta İkizi	Overjet	Ağızdan Solunum	0,374	0,556	0,673	0,501
	Çift Yumurta İkizi	Overjet	Ağızdan Solunum	1,053	0,557	1,891	0,059
Yöntem	**Düzenleyici Grubu**	**Tahmin Edilen Değişken**	**Tahminci Değişken**	**Orijinal Örneklem (Std. β)**	**Std. Sapma**	**t**	**p**
KEKK-ÇGA (Dummy Kodlama İle)	Tek Yumurta İkizi	Overjet	Ağızdan Solunum	0,067	0,102	0,653	0,514
	Çift Yumurta İkizi	Overjet	Ağızdan Solunum	0,163	0,107	1,526	0,127
KEKK-YEM (Nümerik Kodlama İle)		Overjet	Ağızdan Solunum	0,122	0,076	1,598	0,110
		Overjet	Zigosite	-0,018	-0,019	0,267	0,789
		Overjet	Düzenleyici Etki	0,049	0,075	0,647	0,518

Process Macro ve KEKK-ÇGA yöntemleri ile yapılan analizler sonucunda, bu defa tek yumurta ikizlerinde ağızdan solunum ve overjet arasındaki pozitif bir ilişki için birbirlerine çok yakın anlamlılık değeri elde edilmiştir (sırasıyla p=0,50 ve p=0,51). Diğer taraftan çift yumurta ikizleri için Process Macro'nun hesapladığı olasılık değeri 0,11'den 0,06'ya düşerken, KEKK-ÇGA'den elde edilen değer 0,07'den 0,12'ye yükselmiştir. Bu farklılaşma düzenleyici değişken analiz sonuçlarına da etki etmiştir.

Tablo 3.19 Model C2 için düzenleyici değişken analizi sonuçları

Yöntem	Z	p
Hayes Process	0,742	0,389
	Katsayı Farkı	p
KEKK-ÇGA (Dummy Kodlama İle)	-0,096	0,513
	t	p
KEKK-YEM (Nümerik Kodlama İle)	0,647	0,518

Tablo 3.19'da görülebileceği üzere, Process ve KEKK-ÇGA yöntemleri arasında bir önceki simülasyona göre bir farklılaşma mevcuttur. Process için elde edilen düzenleyicilik analizinin sonucu 0,403'den 0,389'a düşerken, KEKK-ÇGA'dan elde edilen değer 0,42'den 0,513'e yükselmiştir. KEKK-YEM'den elde edilen değer ise bu değere çok yakın bir şekilde 0,518'dir.

Aynı model son kez, tasarımı örneklem büyüklüğünün grup başına 30'un üstünde olduğu ve gruplar arasında frekans farkının yüksek olduğu daha büyük bir örneklem ile tekrar çalıştırılmıştır.

3.4.3 Örneklem Büyüklüğünün Grup Başına 30'un Üstünde Olduğu Büyük Örneklem İçin Yöntem Karşılaştırılması

Bu kısımda nominal değişkenlerde deney ve kontrol gruplarının her ikisinin de örneklem büyüklüğünün 30'un üstünde olduğu, toplam örneklem büyüklüğünün ise 404 olduğu daha büyük örneklemli bir çalışma modeli simüle edilmiştir. Bu çalışma için tek yumurta ikizlerinden 118 adet, çift yumurta ikizlerinden ise 286 adet gözlem analize sokulmuştur. Çalışmaya ait örneklemin tanımlayıcı istatistikleri Tablo 3.20'de gösterilmiştir.

Tablo 3.20 Model C3 için tanımlayıcı istatistikler

N=404				
Zigosite	**Değişken**	**Grup**	**n**	**%**
Tek Yumurta İkizi (n=118)	**Ağızdan Soluma**	**Hayır**	76	64,4
		Evet	42	35,6
	Overjet	**Hayır**	91	77,1
		Evet	27	22,9
Çift Yumurta İkizi (n=286)	**Ağızdan Soluma**	**Hayır**	239	83,6
		Evet	47	16,4
	Overjet	**Hayır**	240	83,9
		Evet	46	16,1

3 ayrı yöntemden elde edilen sonuçlar Tablo 3.21'de gösterilmiştir.

Tablo 3.21 Model C3'ten elde edilen katsayılara ait istatistikler

Yöntem	Düzenleyici Grubu	Tahmin Edilen Değişken	Tahminci Değişken	Koşullu Etki	Std. Hata	z	p
Process Macro	Tek Yumurta İkizi	Overjet	Ağızdan Solunum	0,885	0,447	1,979	0,048
	Çift Yumurta İkizi	Overjet	Ağızdan Solunum	1,280	0,365	3,511	0,000
Yöntem	**Düzenleyici Grubu**	**Tahmin Edilen Değişken**	**Tahminci Değişken**	**Orijinal Örneklem (Std. β)**	**Std. Sapma**	**t**	**p**
KEKK-ÇGA (Dummy Kodlama İle)	Tek Yumurta İkizi	Overjet	Ağızdan Solunum	0,185	0,184	0,096	0,054
	Çift Yumurta İkizi	Overjet	Ağızdan Solunum	0,217	0,217	0,072	0,003
KEKK-YEM (Nümerik Kodlama İle)		Overjet	Ağızdan Solunum	0,218	0,06	3,642	0,000
		Overjet	Zigosite	-0,035	0,051	0,677	0,498
		Overjet	Düzenleyici Etki	0,033	0,053	0,616	0,538

Tablo 3.21 incelendiğinde Process Macro ve KEKK-ÇGA yöntemlerinin hem tek yumurta ikizleri hem çift yumurta ikizleri için birbirlerine çok yakın anlamlılık değerleri ürettiği görülmektedir. Fakat bu benzerliğe rağmen her iki yöntemin düzenleyici değişken analizi sonuçları tamamen birbirinden farklılaşmıştır (Tablo 3.22).

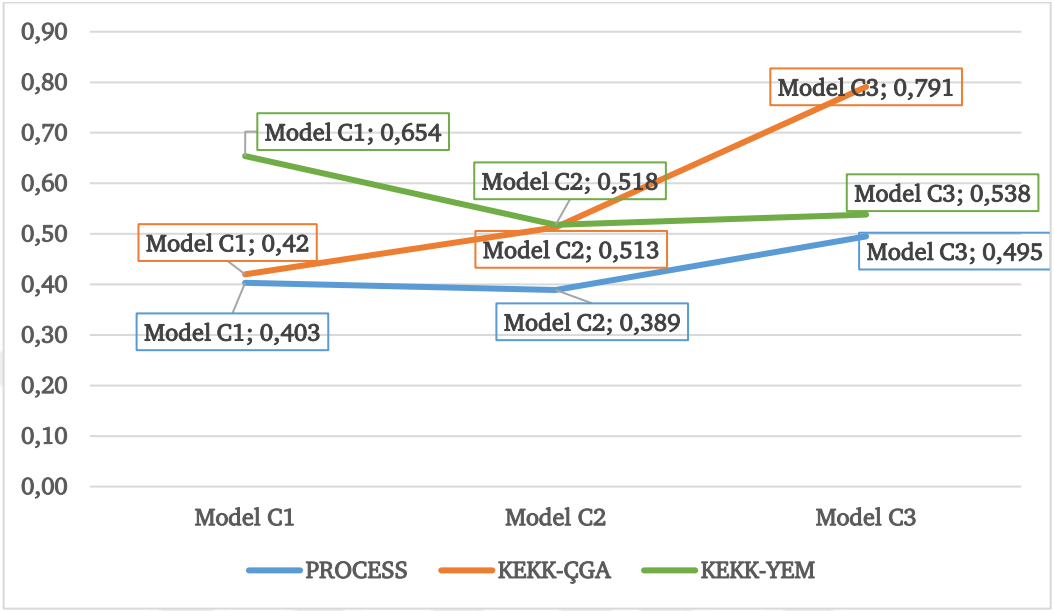
Tablo 3.22 Model C3 için düzenleyici değişken analizi sonuçları

Test	Z	p
Process Macro	-0,466	0,495
	Katsayı Farkı	p
KEKK-ÇGA (Dummy Kodlama İle)	-0,032	0,791
	t	p
KEKK-YEM (Nümerik Kodlama İle)	0,616	0,538

Process Macro yönteminde daha küçük örneklemler ile çalıştırılan C1 ve C2 modelleri için düzenleyici etkiye ait anlamlılık değerleri sırasıyla 0,40 ve 0,39 idi. Bu değer C3 modelinde 0,50 değerine yükseldi.

KEKK-ÇGA yönteminde C1 ve C2 modelleri için düzenleyici etkiye ait anlamlılık değerleri sırasıyla 0,42 ve 0,51 idi. Bu değer C3 modelinde 0,79'a yükseldi.

KEKK-YEM yönteminde C1 ve C2 modelleri için düzenleyici etkiye ait anlamlılık değerleri sırasıyla 0,65 ve 0,51 idi. Bu değer C3 modelinde 0,54 olarak hesaplandı.



Şekil 3.9 3 farklı modelden elde edilen p değerlerinin yönteme göre değişim grafiği

3 ayrı model için 3 farklı yöntemden elde edilen sonuçları birlikte değerlendirildiğinde, öncelikle gruplar arası frekans farkının az olduğu durumlarda, örneklem büyüklükleri 30'un kalsa dahi dummy kodlama kullanılarak KEKK-ÇGA yöntemi tercih edildiğin PROCESS Macro analizi ile benzer sonuçları verdiği söylenebilir. Ancak düzenleyici değişken için gruplar arası frekans farkı arttıkça, KEKK-ÇGA sonuçları PROCESS Macro'ya göre farklılaştığı ve stabil sonuçlar vermediği gözlemlenmiştir. Bu durum KEKK-ÇGA'in gruplar arasındaki asimetrik örneklem büyüklüklerinden kaynaklanabilir.

Diğer taraftan nümerik kodlama kullanılarak düzenleyici değişken modeline KEKK-YEM analizi uygulandığında küçük örneklemlerde PROCESS Macro'ya göre çok daha yüksek olasılık değeri hesaplandığı gözlemlenmiş fakat örneklem büyüklüğü arttırıldıkça iki yöntemin benzer sonuçlar üretmeye başladığı tespit edilmiştir. Nitekim C3 modeli için PROCESS analizinde p=0,538 hesaplanmış iken, KEKK-YEM analizinde p=0,538'dir. Buna göre YEM modellerinde ikili

değişkenlere nümerik kodlama uygulanarak KEKK-YEM tercih edildiğinde farklı örneklem büyüklüklerinde ve farklı frekans dağılımlarında KEKK-ÇGA'ya kıyasla daha kararlı sonuçlar vermesi, gruplar arası örneklem büyüklüğü farkından daha az etkilenmesi beklenebilir.

4
SONUÇ VE ÖNERİLER

Bu tez çalışmasında, KEKK-YEM modellerinde ikili değişkenlerin bağımlı, bağımsız, aracı ve düzenleyici değişken olarak kullanımı farklı model tipleri ve farklı örneklem büyüklerine göre incelenmiştir.

Birinci bölümde, ikili bağımsız değişken ile nümerik bağımlı değişkenin bulunduğu bir model, küçük ve büyük örneklemler ile ayrı ayrı analiz edilmiş, KEKK-YEM ve klasik doğrusal regresyondan elde edilen sonuçlar karşılaştırılmıştır. Yapılan analizler sonucunda örneklem büyüklüğü 30'un altında ve üstünde olduğu durumlar için de ikili değişkenin gruplar arasında frekans çarpıklığı olmadığı sürece her iki yöntemin birbirine eşit ve/veya çok yakın sonuçlar verdiği gözlemlenmiştir. Ancak kontrol veya deney gruplarından biri, diğerinin 7-8 katı kadar daha fazla örnekleme sahip olduğunda yöntemlerden elde edilen sonuçların farklılaşabildiği görülmüştür. Bu durumda, örneklem büyüklüğünün küçük ya da büyük olmasından bağımsız olarak KEKK-YEM'in daha yüksek p olasılık değerleri ürettiği ve örneklem büyüklüğü arttırıldığında her iki yöntemin sonuçlarının yükseldiği görülmüştür. Bu ifadeyi açmak gerekirse, KEKK-YEM'in klasik doğrusal regresyona göre α tipi hataya (1. Tip hataya) ve düşük örneklem büyüklüğüne daha dirençli olması beklenebilir. Bu nedenle araştırmacıların, örneklem büyüklükleri 30'un altında kaldığı ve ikili nominal değişkenleri arasındaki farklar çok yüksek oranda olduğu durumlarda EKK yöntemi ile çalışan klasik doğrusal regresyon yerine KEKK-YEM yöntemini tercih etmeleri, daha güvenilir bir istatistiki analize yardımcı olabilir.

Çalışmanın ikinci kısmında, ikili bağımsız değişken, ikili bağımlı değişken ve ikili aracı değişkenin bulunduğu aracılık modeli araştırılmıştır. Ayrıca bu bölümde, bir ikili değişkenin modelde bağımlı değişken olarak yer alması KEKK-YEM metoduna göre uygun olmadığı için değişkenler hem dummy kodlama ile kullanılmış hem de bir nümerik değişken dönüşümü önerilmiştir. Bu öneriye göre, dummy kodlamada göre 0 ile kodlanan kontrol grup için ortalaması 0, standart sapması olan 1 olan bir standart normal dağılım verisi, 1 ile kodlanan vakaların bulunduğu deney grubu için ise standart sapması 1, ortalaması 10 olan normal dağılım verisi

türetildiğinde KEKK-YEM ile alternatif yöntemlere en yakın sonuçlar elde edilebildiği görülmüştür. Tasarlanan aracı değişken modeline uygulanan analizler sonucunda, örneklem büyüklüğünün grup başına 30'un altında kaldığı küçük örneklemler için dummy kodlamasının yapılması durumunda KEKK-YEM yönteminin lojistik regresyon analizi ile model katsayıları için aynı ve/veya çok yakın p değerlerini ürettiği tespit edilmiştir. Aracılık analizleri incelendiğinde ise KEKK-YEM yönteminin hem dummy hem nümerik kodlama ile uygulandığında, Aroian aracılık testi ile aynı sonucu verdiği gözlemlenmiştir. Sobel testi ise bu 3 yöntemden yaklaşık 0,10 kadar daha düşük bir p değeri hesaplamıştır.

Örneklem büyüklüğünün grup başına 30'un altında kaldığı, diğer bir ifade ile gruplar arası frekans farkının olduğu büyük örneklemler ile çalışıldığında ikili lojistik regresyon ve KEKK-YEM yönteminden elde edilen p değerlerinin farklılaşmaya başladığı, EKK-YEM'in lojistik regresyona kıyasla bazı yol katsayıları için, dummy kodlama kullanıldığında 3 kat, nümerik kodlama ile 2 kat daha yüksek p değerleri tahminlediği görülmüştür. Buna göre grup başına örneklem büyüklüğünün 30'un altında kaldığı ve gruplar arasında frekans farkının yüksek olduğu durumlarda, ikili değişkenleri dummy yerine nümerik kodlama ile kodlamak, araştırmacılara tip 2 hatadan kaçınmak adına yardımcı olabilir. Aracılık analizleri incelendiğinde ise Sobel, Aroian, Goodman testleri ile, KEKK-YEM yönteminin hem dummy hem nümerik kodlama ile uygulandığında dolaylı etkiyi istatistiki olarak yakalayabildiği görülmüştür. Ancak Sobel testlerinin sonuçları %99 güvenle yorumlanabilirken KEKK-YEM analizleri sonuçları %95 güvenle raporlanabilmiştir. Yol katsayılarına ait sonuçların Sobel testleri ile KEKK-YEM arasında farklılaşması, tespit edilen aracılık rolünün yorumlanmasına da etki etmiştir. Sobel testlerine göre kısmi aracılık etkisi, KEKK-YEM için dummy kodlama yapıldığında "tam aracılık", nümerik kodlama yapıldığında "kısmi aracılık" olarak yorumlanabilmiştir.

Grup başına örneklem büyüklüğünün grup başına 30'un üstünde olduğu fakat frekans farkının yine yüksek seviyede olduğu daha büyük bir örneklem ile çalışıldığındaysa kullanılan 5 farklı yöntemin tamamı aynı sonuçları vermiştir. Buna göre grup başına düşen örneklem büyüklüklerinin 30'dan yüksek olduğu durumlarda, gruplar arası frekans farkı çok yüksek olmasına rağmen, KEKK-

YEM'in hem dummy hem de nümerik kodlama tercih edildiğinde Sobel testleri ile aynı sonuçlar vererek modeldeki dolaylı etkiyi tespit edebildiği söylenebilir.

Bu noktada değinilmesi gereken önemli bir husus, aracılık testi için kullanılabilecek yöntemlerin kısıtlamalarıdır. Nitekim Sobel, Aroian ve Goodman testleri sadece bir bağımsız, bir aracı ve bir nümerik değişken olmak üzere 3 değişkenli modeller için kullanılabilirken, PROCESS Macro, modelde aracı değişken olarak kullanılan değişken sadece nümerik olduğunda hesaplama yapmaktadır. Dolayısıyla, 2 veya daha fazla sayıda kategorik aracı kullanarak bir aracılık modellemesi çalışmak isteyen araştırmacıların tercih edebilecekleri tek tercih olarak KEKK-YEM yöntemi bulunmaktadır. Fakat bu yöntemi kullanırken araştırmacıların değişken kodlamalarını dikkatle yapmalarında fayda vardır. Nitekim grupların örneklem büyüklükleri eşit ya da yakın olduğunda dummy kodlama tutarlı sonuçlar verebilmektedir. Fakat deney ya da kontrol gruplarına ait örneklem büyüklüğü 30'un altında kaldığında ve gruplar arasında frekans farkı çok yüksek olduğunda KEKK-YEM yöntemi için nümerik kodlama dummy kodlamaya kıyasla daha güçlü sonuçlar verebilmektedir.

Ayrıca küçük örneklemlerle çalışıldığında, eğer dummy kodlama tercih edilirse KEKK-YEM algoritması 5000 ya da 10000 yeniden örnekleme ile sonuçlar verememektedir. Yeniden örnekleme sayısı 1000 kez yapıldığında ise kimi zaman bazı yol katsayılarının, kimi zamansa hiçbir yol katsayısının hesaplanamadığı görülmüştür. Bu durumlar SmartPLS programı kullanıcıya çözüm yolu olarak örneklem büyüklüğünün genişletilmesini önermektedir. Nitekim bütün grupların örneklem büyüklükleri 30'un üstünde olduğu büyük örneklemli denemelerde yeniden örnekleme sayısı 5000'e çıkarılabilmiştir.

Ancak özellikle laboratuvar ortamında yürütülen bilimsel çalışmalarda, etik ve maddi sebeplerle örneklem büyüklüğünü arttırmak her zaman mümkün olmayabilmektedir. Bu çözüme bir alternatif olarak, dummy yerine nümerik kodlama yapılması, hem daha çok sayıda yeniden örnekleme yapılmasını hem de yapılan testlerin istatistiksel gücünün daha yüksek seviyede olmasını sağlayabilir.

Çalışmanın üçüncü kısmında, ikili bağımsız değişken, ikili bağımlı değişken ve ikili düzenleyici değişkenin bulunduğu düzenleyici değişken modeli araştırılmıştır.

Sonuçların PROCESS Macro, KEKK-ÇGA ve KEKK-YEM olmak üzere 3 farklı yöntem kullanılarak elde edildiği dikkate alındığında, ilk olarak gruplar arasındaki frekans farkının düşük olduğu durumlarda örneklem büyüklükleri 30 ve altında olmasına rağmen KEKK-ÇGA yöntemi PROCESS Macro analiziyle benzer sonuçlar verdiği gözlenmiştir. Ancak düzenleyici değişken için gruplar arasındaki frekans farkı arttıkça, KEKK-ÇGA sonuçlarının PROCESS Macro'ya göre farklılaştığı ve istikrarlı sonuçlar sağlamadığı tespit edilmiştir. Bunun sebebi, KEKK-ÇGA'nın gruplar arasındaki asimetrik örneklem büyüklüklerine dirençsiz olması olabilir.

Öte yandan, düzenleyici değişken modeline nümerik kodlama kullanılarak KEKK-YEM analizi uygulandığında, küçük örneklemlerde düzenleyicilik etkisi için PROCESS Macro'ya kıyasla çok daha yüksek olasılık değerleri hesaplandığı görülmüş ancak örneklem büyüklüğü arttıkça iki yöntemin benzer sonuçlar üretmeye başladığı tespit edilmiştir. Bu nedenle, YEM modellerinde nümerik kodlama ve KEKK-YEM yönteminin tercih edilmesi, farklı örneklem büyüklüklerinde ve farklı frekans dağılımlarında, KEKK-ÇGA'ya kıyasla daha kararlı sonuçlar elde edilmesine katkıda bulunabilir.

KAYNAKÇA

[1] R. Richardo *et al.*, "The impact of STEM attitudes and computational thinking on 21st-century via structural equation modelling," *International Journal of Evaluation and Research in Education (IJERE),* 2023.

[2] V. Shanmugam and J. E. Marsh, "Application of Structural Equation Modeling to the Social Sciences: A Brief Guide for Researchers," 2016.

[3] M. Matande, A. Thoyib, and D. T. Kurniawati, "The effect of perceived organizational support and transformational leadership on turnover intention of health workers at Hospital X Mimika Regency Papua," *International Journal of Research in Business and Social Science (2147-4478),* 2022.

[4] U. Wahyuni and D. Permana, "Consumer Behavior through Online Shop : Continuance Intention Theory Approach," *International Journal of Scientific Research in Science and Technology,* 2022.

[5] K. K.-K. Wong, *Mastering partial least squares structural equation modeling (PLS-Sem) with Smartpls in 38 Hours*. IUniverse, 2019.

[6] C. Fornell and F. L. Bookstein, "Two structural equation models: LISREL and PLS applied to consumer exit-voice theory," *Journal of Marketing research,* vol. 19, no. 4, pp. 440-452, 1982.

[7] M. Haenlein and A. M. Kaplan, "A beginner's guide to partial least squares analysis', Understanding Stafisfics 3 (4), 283–297," *Google Scholar Google Scholar Cross Ref Cross Ref,* 2004.

[8] W. Reinartz, M. Haenlein, and J. Henseler, "An empirical comparison of the efficacy of covariance-based and variance-based SEM," *International Journal of research in Marketing,* vol. 26, no. 4, pp. 332-344, 2009.

[9] J. F. Hair, M. Sarstedt, C. M. Ringle, and J. A. Mena, "An assessment of the use of partial least squares structural equation modeling in marketing research," *Journal of the academy of marketing science,* vol. 40, pp. 414-433, 2012.

[10] S. Ekemhonye, J. Nmadu, A. Coker, and M. Ndanitsa, "The Determinants of Farmers Adaptive Capacities to Poverty Related Diseases along River Niger in Edo and Kogi States, Nigeria," *Asian Journal of Advances in Agricultural Research,* vol. 21, no. 4, pp. 39-45, 2023.

[11] M. J. Retutas and M. T. Rubio, "Multivariate Analysis on Performance in Statistics, Self-Efficacy and Attitudes of Senior High School Students," *Journal of Research and Advances in Mathematics Education,* vol. 6, no. 4, pp. 352-367, 2021.

[12] V. Shanmugam and J. E. Marsh, "Application of structural equation modeling to the social sciences: A brief guide for researchers," *Mesure et évaluation en éducation,* vol. 37, no. 3, pp. 99-123, 2015.

[13] J. F. Hair, W. Black, B. Babin, and R. Anderson, "Multivariate data analysis (8. Baskı)," *Eight Edition, Cengage: Learning EMEA,* 2019.

[14] C. ALPAR, "Uygulamalı çok değişkenli istatistiksel yöntemler," 2017.

[15] A. Arzu and H. Onder, "Farkli veri yapilarinda kullanilabilecek regresyon yöntemleri," *Anadolu Tarım Bilimleri Dergisi,* vol. 28, no. 3, pp. 168-174, 2013.

[16] Ö. Yeniay and A. GÖKTAŞ, "A comparison of partial least squares regression with other prediction methods," *Hacettepe Journal of Mathematics and Statistics,* vol. 31, pp. 99-111, 2002.

[17] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*. John Wiley & Sons, 2013.

[18] D. G. Kleinbaum and M. Klein, "Analysis of matched data using logistic regression," *Logistic regression: A self-learning text,* pp. 227-265, 2002.

[19] A. G. E. Polat and S. Günay, "Kismi en küçük kareler ve bir uygulama," *Içinde VI: İstatistik Günleri Sempozyumu Bildiriler Kitabı,* 2009.

[20] H. Abdi, "Partial Least Square Regression: PLS Regression," *Encyclopedia of Measurementand Statistics. Thousand Oaks (CA):Sage.,* pp. 1-13, 2007.

[21] H. Wold, "Estimation of principal components and related models by iterative least squares," *Multivariate analysis,* pp. 391-420, 1966.

[22] S. Wold, M. Sjostrom, and L. Eriksson, "PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratoary Systems 58: 109–130," ed, 2001.

[23] P. H. Garthwaite, "An interpretation of partial least squares," *Journal of the American Statistical Association,* vol. 89, no. 425, pp. 122-127, 1994.

[24] S. De Jong, "SIMPLS: an alternative approach to partial least squares regression," *Chemometrics and intelligent laboratory systems,* vol. 18, no. 3, pp. 251-263, 1993.

[25] S. Wold, A. Ruhe, H. Wold, and I. Dunn, WJ, "The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses," *SIAM Journal on Scientific and Statistical Computing,* vol. 5, no. 3, pp. 735-743, 1984.

[26] D. W. Hosmer and S. Lemesbow, "Goodness of fit tests for the multiple logistic regression model," *Communications in statistics-Theory and Methods,* vol. 9, no. 10, pp. 1043-1069, 1980.

[27] A. Agresti, "An introduction to categorical data analysis," ed: Taylor & Francis, 1996.

[28] H. Tatlidil, "Uygulamali çok degiskenli istatistiksel analiz," *Ankara: Cem Web Ofset Ltd. Sti,* 1996.

[29] B. G. Tabachnick, L. S. Fidell, and J. B. Ullman, *Using multivariate statistics*. pearson Boston, MA, 2013.

[30] R. E. Schumacker and R. G. Lomax, *A beginner's guide to structural equation modeling*. psychology press, 2004.

[31] S. Wright, "Correlation and causation," 1921.

[32] S. Wright, "The method of path coefficients," *The annals of mathematical statistics,* vol. 5, no. 3, pp. 161-215, 1934.

[33] S. Wright, "Path coefficients and path regressions: alternative or complementary concepts?," *Biometrics,* vol. 16, no. 2, pp. 189-202, 1960.

[34] C. Tilly, "Big structures, large processes, huge comparisons," 1984.

[35] T. Haavelmo, "The statistical implications of a system of simultaneous equations," *Econometrica, Journal of the Econometric Society,* pp. 1-12, 1943.

[36] K. A. Bollen, *Structural equations with latent variables*. John Wiley & Sons, 1989.

[37] A. S. Goldberger and I. Olkin, "A minimum-distance interpretation of limited-information estimation," *Econometrica,* vol. 39, 1971// 1971.

[38] A. S. Goldberger, "Structural Equation Methods in the Social Sciences," *Econometrica,* vol. 40, no. 6, pp. 979-1001, 1972, doi: 10.2307/1913851.

[39] D. Wiley, A. Goldberger, and O. Duncan, "Structural equation models in the social sciences," *AS Goldberger & OD Duncan (Eds.),* pp. 69-83, 1973.

[40] P. Tarka, "An overview of structural equation modeling: its beginnings, historical development, usefulness and controversies in the social sciences," *Quality & quantity,* vol. 52, pp. 313-354, 2018.

[41] K. G. Jöreskog, "A general method for analysis of covariance structures," *Biometrika,* vol. 57, 1970// 1970.

[42] K. Joreskog, "A general method for estimating a linear structural equation system In Goldberger AS, & Duncan OD (Eds.), Structural equation models in the social sciences," 1973: New York, NY: Seminar Press.[Google Scholar].

[43] J. W. Keesling, "MAXIMUM LIKELIHOOD APPROACHES TO CAUSAL FLOW ANALYSIS," The University of Chicago, 1973.

[44] K. G. Jöreskog and D. Sörbom, *LISREL IV: A general computer program for estimation of linear structural equation systems by maximum likelihood methods*. University of Uppsala, Department of statistics [Uppsala univ., Statistiska …, 1978.

[45] R. M. Hauser and A. S. Goldberger, "The treatment of unobservable variables in path analysis," *Sociological methodology,* vol. 3, pp. 81-117, 1971.

[46] P. M. Bentler, *Theory and implementation of EQS: A structural equations program*. BMDP Statistical Software, 1985.

[47] B. O. Muthen, "LISCOMP: Analysis of linear structural equations with a comprehensive measurement model," *Mooresville, IN: Scientific Software,* 1987.

[48] L. Muthén and B. Muthén, "Mplus user's guide: Statistical analysis with latent variables (1998-2012)," *Los Angeles, CA,* 1998.

[49] J. L. Arbuckle and W. Wothke, *Amos 4.0 user's guide*. Citeseer, 1999.

[50] A. Bryk, S. Raudenbush, and R. Congdon, "HLM 4 for Windows (Version 6.08)[Computer software]. Chicago, IL: Scientific Software International, Inc," ed, 1996.

[51] K. G. Jöreskog and D. Sörbom, *LISREL 8: User's reference guide*. Scientific Software International, 1996.

[52] S. Rabe-Hesketh, A. Skrondal, and A. Pickles, "Generalized multilevel structural equation modeling," *Psychometrika,* vol. 69, no. 2, pp. 167-190, 2004.

[53] A. Skrondal and S. Rabe-Hesketh, *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Crc Press, 2004.

[54] J. Fox, "Teacher's corner: structural equation modeling with the sem package in R," *Structural equation modeling,* vol. 13, no. 3, pp. 465-486, 2006.

[55] Y. Rosseel, "lavaan: An R package for structural equation modeling," *Journal of statistical software,* vol. 48, pp. 1-36, 2012.

[56] J. C. Westland, "Structural equation models," *Stud. Syst. Decis. Control,* vol. 22, no. 5, p. 152, 2015.

[57] S. L. Hershberger, "The growth of structural equation modeling: 1994-2001," *Structural Equation Modeling,* vol. 10, no. 1, pp. 35-46, 2003.

[58] J. Hair, C. Ringle, and M. Sarstedt, "PLS-SEM: Indeed a Silver Bullet. Journal of Marketing Theory and Practice," *Journal of Marketing Theory and Practice,* vol. 19, no. 2, 2015.

[59] H. Wold, "Path models with latent variables: The NIPALS approach," in *Quantitative sociology*: Elsevier, 1975, pp. 307-357.

[60] M. TENENHAUS, Y. C. Vinzia, and C. Lauro, "PLS Path Modeling Computational Statistics & Data Analysis, v. 48," 2005.

[61] J. F. Hair, C. M. Ringle, and M. Sarstedt, "PLSSEM: Indeed a Silver Bullet Journal of Marketing Theory and Practice, 19 (2) pages 139-152, doi: 10.2753," *MTP1069-6679190202,* 2011.

[62] H. Hwang, N. K. Malhotra, Y. Kim, M. A. Tomiuk, and S. Hong, "A comparative study on parameter recovery of three approaches to structural equation modeling," *Journal of Marketing research,* vol. 47, no. 4, pp. 699-712, 2010.

[63] R. B. Kline, *Principles and practice of structural equation modeling*. Guilford publications, 2023.

[64] L. t. Hu and P. M. Bentler, "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," *Structural equation modeling: a multidisciplinary journal,* vol. 6, no. 1, pp. 1-55, 1999.

[65] J. B. Schreiber, A. Nora, F. K. Stage, E. A. Barlow, and J. King, "Reporting structural equation modeling and confirmatory factor analysis results: A review," *The Journal of educational research,* vol. 99, no. 6, pp. 323-338, 2006.

[66] T. D. Little, *Longitudinal structural equation modeling*. Guilford press, 2013.

[67] B. M. Byrne, *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. routledge, 2013.

[68] H. W. Marsh, O. Lüdtke, U. Trautwein, and A. J. Morin, "Classical latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered approaches to theoretical models of self-concept," *Structural Equation Modeling: A Multidisciplinary Journal,* vol. 16, no. 2, pp. 191-225, 2009.

[69] S. Wasserman and K. Faust, "Social network analysis: Methods and applications," 1994.

[70] W. W. Chin and P. R. Newsted, "Structural equation modeling analysis with small samples using partial least squares," *Statistical strategies for small sample research,* vol. 1, no. 1, pp. 307-341, 1999.

[71] B. Muthén and T. Asparouhov, "Bayesian structural equation modeling: a more flexible representation of substantive theory," *Psychological methods,* vol. 17, no. 3, p. 313, 2012.

[72] C.-Y. Yu, *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*. University of California, Los Angeles, 2002.

[73] R. J. Tibshirani and B. Efron, "An introduction to the bootstrap," *Monographs on statistics and applied probability,* vol. 57, no. 1, 1993.

[74] J. R. Quinlan, "Induction of decision trees," *Machine learning,* vol. 1, pp. 81-106, 1986.

[75] M. Stone, "Cross-validatory choice and assessment of statistical predictions," *Journal of the royal statistical society: Series B (Methodological),* vol. 36, no. 2, pp. 111-133, 1974.

[76] L. D. Bacon, "Using LISREL and PLS to measure customer satisfaction," in *Sawtooth Software Conference Proceedings*, 1999: La Jolla California, pp. 2-5.

[77] K. Wong, "Handling small survey sample size and skewed dataset with partial least square path modelling," *Vue: the magazine of the Marketing Research and Intelligence Association,* vol. 20, pp. 20-23, 2010.

[78] J. F. Hair Jr, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications, 2021.

[79] H. Olya, "Partial least squares based structural equation modeling (PLS-SEM)," in *Global Conference on Services Management*, 2017, vol. 3, no. 7.

[80] B. M. Byrne, "Structural equation modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument," *International journal of testing,* vol. 1, no. 1, pp. 55-86, 2001.

[81] C. K. Enders, *Applied missing data analysis*. Guilford Publications, 2022.

[82] R. P. Bagozzi and Y. Yi, "Specification, evaluation, and interpretation of structural equation models," *Journal of the academy of marketing science,* vol. 40, pp. 8-34, 2012.

[83] A. Bryman, *Social research methods*. Oxford university press, 2016.

[84] B. Tabachnick and L. Fidell, "Using Multivariate Statistics, ; Ullman, JB, Ed," ed: Pearson: Boston, MA, USA, 2019.

[85] J. F. Hair Jr, M. Sarstedt, C. M. Ringle, and S. P. Gudergan, *Advanced issues in partial least squares structural equation modeling*. saGe publications, 2017.

[86] J. Henseler, C. M. Ringle, and M. Sarstedt, "A new criterion for assessing discriminant validity in variance-based structural equation modeling," *Journal of the academy of marketing science,* vol. 43, pp. 115-135, 2015.

[87] W. W. Chin, "The partial least squares approach to structural equation modeling," *Modern methods for business research,* vol. 295, no. 2, pp. 295-336, 1998.

[88] A. Diamantopoulos and H. M. Winklhofer, "Index construction with formative indicators: An alternative to scale development," *Journal of marketing research,* vol. 38, no. 2, pp. 269-277, 2001.

[89] J. Hulland, "Use of partial least squares (PLS) in strategic management research: A review of four recent studies," *Strategic management journal,* vol. 20, no. 2, pp. 195-204, 1999.

[90] B. M. Byrne, *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. routledge, 2016.

[91] J.-B. Lohmöller, *Latent variable path modeling with partial least squares*. Springer Science & Business Media, 1989.

[92] G. Lamberti, "Modeling with Heterogeneity," Doctoral Dissertation, Universitat Politecnica De Catalunya, Barcelona, 2015.

[93] V. E. Vinzi, L. Trinchera, and S. Amato, "PLS path modeling: from foundations to recent developments and open issues for model assessment and improvement," *Handbook of partial least squares: Concepts, methods and applications,* pp. 47-82, 2010.

[94] A. Crisci, "ESTIMATION METHODS FOR THE STRUCTURAL EQUATION MODELS: MAXIMUM LIKELIHOOD, PARTIAL LEAST SQUARES E GENERALIZED MAXIMUM ENTROPY," *Journal of applied quantitative methods,* vol. 7, no. 2, 2012.

[95] Y. Kinaş, "Kısmi En Küçük Kareler Yapısal Eşitlik Modellemesiyle Aracılık ve Düzenleyicilik Etkilerinin Araştırılması ve Bir Uygulama," Doktora Tezi, Eskişehir Osmangazi Üniversitesi, 2021.

[96] H. Wold, "Partial Least Squares. Encyclopedia of statistical sciences. vol 6 Kotz, S & Johnson, NL," ed: John Wiley and sons, New York, 1985.

[97] H. Wold, "Soft modelling: the basic design and some extensions," *Systems under indirect observation, Part II,* pp. 36-37, 1982.

[98] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *psychometrika,* vol. 16, no. 3, pp. 297-334, 1951.

[99] J. C. Nunnally, V. Arellano, and J. Albertotr, *Teoría psicométrica* (no. 152.8 N8). 1995.

[100] J. F. Hair, C. M. Ringle, and M. Sarstedt, "PLS-SEM: Indeed a Silver Bullet," *Journal of Marketing Theory and Practice,* vol. 19, no. 2, pp. 139-152, 2011/04/01 2011, doi: 10.2753/MTP1069-6679190202.

[101] N. Urbach and F. Ahlemann, "Structural equation modeling in information systems research using partial least squares," *Journal of Information Technology Theory and Application (JITTA),* vol. 11, no. 2, p. 2, 2010.

[102] J. Henseler, G. Hubona, and P. A. Ray, "Using PLS path modeling in new technology research: updated guidelines," *Industrial management & data systems,* 2016.

[103] D. Doğan, "SmartPLS ile veri analizi," *Ankara: Zet Yayınları,* vol. 2, 2019.

[104] R. P. Bagozzi and Y. Yi, "On the evaluation of structural equation models," *Journal of the academy of marketing science,* vol. 16, pp. 74-94, 1988.

[105] W. Afthanorhan, "A comparison of partial least square structural equation modeling (PLS-SEM) and covariance based structural equation modeling (CB-SEM) for confirmatory factor analysis," *International Journal of Engineering Science and Innovative Technology,* vol. 2, no. 5, pp. 198-205, 2013.

[106] C. Fornell and D. F. Larcker, "Evaluating structural equation models with unobservable variables and measurement error," *Journal of marketing research,* vol. 18, no. 1, pp. 39-50, 1981.

[107] J. F. Hair, R. E. Anderson, R. L. Tatham, and W. C. Black, "Multivariate data analysis. englewood cliff," *New jersey, USA,* vol. 5, no. 3, pp. 207-2019, 1998.

[108] W. W. Chin, B. L. Marcolin, and P. R. Newsted, "A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail

emotion/adoption study," *Information systems research,* vol. 14, no. 2, pp. 189-217, 2003.

[109] K. K.-K. Wong, "Partial least squares structural equation modeling (PLS-SEM) techniques using SmartPLS," *Marketing Bulletin,* vol. 24, no. 1, pp. 1-32, 2013.

[110] P. M. Bentler and D. G. Bonett, "Significance tests and goodness of fit in the analysis of covariance structures," *Psychological bulletin,* vol. 88, no. 3, p. 588, 1980.

[111] J. Henseler, "Partial least squares path modeling: Quo vadis?," *Quality & Quantity,* vol. 52, no. 1, pp. 1-8, 2018.

[112] J. Henseler, C. M. Ringle, and M. Sarstedt, "Testing measurement invariance of composites using partial least squares," *International marketing review,* 2016.

[113] J. Cohen, *Statistical power analysis for the behavioral sciences*. Academic press, 2013.

[114] R. M. Baron and D. A. Kenny, "The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations," *Journal of personality and social psychology,* vol. 51, no. 6, p. 1173, 1986.

[115] D. P. MacKinnon, A. J. Fairchild, and M. S. Fritz, "Mediation analysis," *Annu. Rev. Psychol.,* vol. 58, pp. 593-614, 2007.

[116] A. F. Hayes, *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications, 2017.

[117] A. F. Hayes and K. J. Preacher, "Conditional process modeling: Using structural equation modeling to examine contingent causal processes," 2013.

[118] K. A. Bollen and J. Pearl, "Eight myths about causality and structural equation models," *Handbook of causal analysis for social research,* pp. 301-328, 2013.

[119] T. J. VanderWeele, "Mediation analysis: a practitioner's guide," *Annual review of public health,* vol. 37, pp. 17-32, 2016.

[120] K. J. Preacher and A. F. Hayes, "Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models," *Behavior Research Methods,* vol. 40, no. 3, pp. 879-891, 2008/08/01 2008, doi: 10.3758/BRM.40.3.879.

[121] M. E. Sobel, "Asymptotic confidence intervals for indirect effects in structural equation models," *Sociological methodology,* vol. 13, pp. 290-312, 1982.

[122] D. P. MacKinnon, C. M. Lockwood, J. M. Hoffman, S. G. West, and V. Sheets, "A comparison of methods to test mediation and other intervening variable effects," *Psychological methods,* vol. 7, no. 1, p. 83, 2002.

[123] J. R. Edwards and L. S. Lambert, "Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis," *Psychological methods,* vol. 12 1, pp. 1-22, 2007.

[124] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken, *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge, 2013.

[125] M. Rönkkö, C. N. McIntosh, and J. Antonakis, "On the adoption of partial least squares in psychological research: Caveat emptor," *Personality and Individual Differences,* vol. 87, pp. 76-84, 2015.

[126] J. Henseler, C. M. Ringle, and R. R. Sinkovics, "The use of partial least squares path modeling in international marketing," in *New challenges to international marketing*: Emerald Group Publishing Limited, 2009.

[127] A. F. Hayes, *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications, 2013.

[128] B. L. De Stavola, R. M. Daniel, G. B. Ploubidis, and N. Micali, "Mediation analysis with intermediate confounding: structural equation modeling viewed through the causal inference lens," *American journal of epidemiology,* vol. 181, no. 1, pp. 64-80, 2015.

[129] J. Fox, "Effect analysis in structural equation models: Extensions and simplified methods of computation," *Sociological Methods & Research,* vol. 9, no. 1, pp. 3-28, 1980.

[130] R. L. Holbert and M. T. Stephenson, "The Importance of Indirect Effects in Media Effects Research: Testing for Mediation in Structural Equation Modeling," *Journal of Broadcasting & Electronic Media,* vol. 47, pp. 556 - 572, 2003.

[131] S. Birant, M. Veznikli, Y. Kasimoglu, M. Koruyucu, A. A. Evren, and F. Seymen, "Path Analysis of the Relationships between the Eruption Time of the First Primary Teeth and Various Factors in Twins," *Children,* vol. 10, no. 4, p. 683, 2023.

[132] *IBM SPSS Statistics for Windows*. (2017). Armonk, NY: IBM Corp.

[133] *Ringle, C.M., Wende, S. and Becker, J.M SmartPLS 3*. (2015). Boenningstedt: SmartPLS GmbH. [Online]. Available: http://www.smartpls.com

[134] *K. J. Preacher ve G. J. Leonardelli, "Calculation for the Sobel Test: An Interactive Calculation Tool for Mediation Tests" [Çevrimiçi Yazılım]*. (2010). [Online]. Available: https://quantpsy.org/sobel/sobel.htm

TEZDEN ÜRETİLMİŞ YAYINLAR

Konferans Bildirileri

1. Veznikli M., Birant S., Koruyucu M., Kasimoglu Y., Evren A. A., Seymen F. “Associations of Genetic, Maternal, Perinatal and Postnatal Factors with the Eruption Timing of The First Primary Tooth in Twins” 4. International Dicle Scientific Research and Innovation Congress 18-19 April 2023. Diyarbakır

Makaleler

1. Birant, S., Veznikli, M., Kasimoglu, Y., Koruyucu, M., Evren, A. A., & Seymen, F. (2023). Path Analysis of the Relationships between the Eruption Time of the First Primary Teeth and Various Factors in Twins. Children, 10(4), 683.