T. C.

ORDU ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

KATEGORİK DEĞİŞKENLER ARASI İLİŞKİ KATSAYILARININ SİMÜLASYON YOLUYLA KARŞILAŞTIRILMASI

SİNEM ŞENSOY

YÜKSEK LİSANS TEZİ

ZOOTEKNİ ANABİLİM DALI

ORDU 2020

TEZ BİLDİRİMİ

Tez yazım kurallarına uygun olarak hazırlanan ve kullanılan intihal tespit programının sonuçlarına göre; bu tezin yazılmasında bilimsel ahlak kurallarına uyulduğunu, başkalarının eserlerinden yararlanılması durumunda bilimsel normlara uygun olarak atıfta bulunulduğunu, tezin içerdiği yenilik ve sonuçların başka bir yerden alınmadığını, kullanılan verilerde herhangi bir tahrifat yapılmadığını, tezin herhangi bir kısmının bu üniversite veya başka bir üniversitedeki başka bir tez çalışması olarak sunulmadığını beyan ederim.

SİNEM ŞENSOY

Not: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

ÖZET

KATEGORİK DEĞİŞKENLER ARASI İLİŞKİ KATSAYILARININ SİMÜLASYON YOLUYLA KARŞILAŞTIRILMASI

SİNEM ŞENSOY

ORDU ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

ZOOTEKNİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ, 53 SAYFA

(TEZ DANIŞMANI: DR. ÖĞR. ÜYESİ YELİZ KAŞKO ARICI)

Bu tez çalışmasının amacı, kategorik değişkenler arasındaki ilişkilerin belirlenmesinde kullanılan ilişki katsayılarının incelenmesi ve bazı ilişki katsayılarının belirlenen deneme koşullarındaki performanslarının karşılaştırılmasıdır. Bu amaçla bir simülasyon çalışması planlanmış ve aralarında sırasıyla 0.5 ve 0.9 korelasyon bulunan iki değişkenli standart normal dağılımdan tesadüf sayıları üretilmiştir. Örneklem genişlikleri sırasıyla 30, 50, 100, 150 ve 200 olarak belirlenmiştir. Tesadüf sayıları eşit aralıklı olarak bölünmüş ve 3×3, 4×4 ve 5×5 ölçekli tablo boyutu olacak şekilde kodlanmıştır. Farklı tablo boyutu, ilişki düzeyi ve örneklem genişliklerinde Pearson, Spearman rank, Kendall tau-b, Kendall tau-c, Goodman Kruskal'ın Gamma ve Somer'in d katsayılarının performansları karşılaştırılmıştır.

Yapılan araştırmalarda genellikle birden fazla değişken ile çalışılır. Bunun sebebi üzerinde çalışılan değişkenin aynı zamanda başka değişkenlerin etkisi altında değişim gösterebilmesidir. Dolayısıyla değişkenler arasındaki ilişkilerin ortaya çıkarılması gerekmektedir. Değişkenler arasındaki ilişkilerin ortaya çıkarılması amacıyla ilişki (korelasyon) katsayıları hesaplanmaktadır. İlişki katsayıları hem değişkenler arasındaki ilişkinin derecesini hem de bu ilişkinin yönünü ortaya koymaktadır. Değişken tipine, dağılımın şekline ve örneklemin genişliğine uygun olarak hesaplanması önerilen farklı ilişki katsayıları geliştirilmiştir. Kategorik tipteki değişkenler için kullanılacak ilişki katsayısı kategori sayısına ve kategoriler arasında sıralama olup olmamasına göre de değişim göstermektedir. Örneğin, isimsel (nominal) değişkenlerde; iki kategoriliier için Phi katsayısı, ikiden fazla kategoriliier için ise Cramer'in V ve Goodman Kruskal'ın Lambda katsayısı kullanılmaktadır. Sıralı (ordinal) değişkenler arasındaki ilişkiler ise duruma göre Spearman rank ilişki katsayısı, Kendall tau-b, Kendall tau-c veya Somer'in d katsayısı ile belirlenmektedir. Bu ilişki katsayıları arasında hangisinin tercih edilmesi gerektiğine karar vermek araştırıcılar için problem olabilmektedir.

Anahtar Kelimeler: Kategorik Değişken, Çapraz Tablo, İlişki Katsayısı, Simülasyon, Sıralı Değişken.

ABSTRACT

A COMPARISON VIA SIMULATION OF THE COEFFICIENTS OF ASSOCIATION BETWEEN CATEGORICAL VARIABLES

SİNEM ŞENSOY

ORDU UNIVERSITY INSTITUTE OF NATURAL AND APPLIED SCIENCES

ANIMAL SCIENCIE

MASTER'S THESIS, 53 PAGES

(SUPERVISOR: ASSIST. PROF. DR. YELİZ KAŞKO ARICI)

The purpose of this thesis study is to examine the relationship coefficients used in determining the relationships between categorical variables and to compare the performance of some relationship coefficients in the determined trial conditions. For this purpose, a simulation study was planned, and random numbers were generated from the standard normal distribution with two variables, with a correlation of 0.5 and 0.9 respectively. Sample widths were determined as 30, 50, 100, 150 and 200 respectively. The coincidence numbers are divided equally and the table is coded to be 3×3, 4×4 and 5×5 scales. The performances of Pearson, Spearman rank, Kendall tau-b, Kendall tau-c, Goodman Kruskal's Gamma and Somer's d coefficients were compared in different table sizes, relationship levels and sample widths.

In researches, it is generally worked with more than one variable. The reason for this is that the variable studied can also change under the influence of other variables. Therefore, the relationships between the variables need to be revealed. Relationship (correlation) coefficients are calculated to reveal the relationships between the variables. Relationship coefficients reveal both the degree of the relationship between the variables and the direction of this relationship. Different relation coefficients, which are recommended to be calculated in accordance with the type of variable, the shape of the distribution and the width of the sample, have been developed. The relation coefficient to be used for categorical variables varies according to the number of categories and whether there is a ranking among the categories. For example; In nominal variables, Phi coefficient is used for two categories, Cramer's V or Goodman Kruskal's Lambda coefficients are used for more than two categories, whereas ordinal variables use Spearman rank relation coefficient, Kendall tau-b, Kendall tau-c or Somer's d coefficient. Deciding which one to choose among these relationship coefficients can be a problem for researchers.

Keywords: Association Coefficient, Categorical Data, Cross Table, Simulation, Ordinal Data.

TEŞEKKÜR

Tez konumun belirlenmesi, çalışmanın yürütülmesi ve yazımı esnasında bilgi ve deneyimlerinden yararlandığım, bana zamanını ayırıp ilgiyle elinden geleni fazlasıyla sunan, tezimi baştan sona kadar okuyup eksiklerini tamamlayan ve söylediği her kelimenin önemini iyi bildiğim kıymetli danışman hocam Sayın Dr. Öğr. Üyesi Yeliz KAŞKO ARICI'ya, tez çalışmam süresince desteğini esirgemeyen, tavsiyeleri ve olumlu yaklaşımlarıyla pozitif düşünmeye yönlendiren değerli hocam Sayın Prof. Dr. Sezai ALKAN'a, öneri ve eleştirileriyle tezimin değerlendirilmesinde yer alan ve önemli katkılar sağlayan jüri üyesi Sayın Dr. Öğr. Üyesi Yasemin GEDİK'e sonsuz teşekkürlerimi sunarım.

Aynı zamanda, bütün hayatım boyunca manevi ve maddi beni destekleyip cesaretlendiren, her daim yanımda olan babam Nazım ŞENSOY ve annem Hüsne ŞENSOY'a, ayrıca hayatımın her evresinde bana yapabileceklerimin sınırı olmadığını hatırlatarak yanımda olan değerli arkadaşlarım Nursaç TEKSAYTAŞ ve Tuğba KESKİN'e çok teşekkür ederim.

İÇİNDEKİLER

Sayfa

ŞEKİL LİSTESİ

ÇİZELGE LİSTESİ

Sayfa

SİMGELER ve KISALTMALAR LİSTESİ

CC	:	Contingency Katsayısı
C_{max}	:	En Büyük Toplam Sütun Gözlem Sayısı
c	:	Sütun Sayısı
d	:	Somer'in d Katsayısının En Yüksek Olabilirlik Tahmini
D_i	:	$X_i - Y_i$ olup, X_i ve Y_i Çiftlerinin Sıra Numaraları Arasındaki Fark
$\hat{d}_{YX}$	:	X Bağımsız, Y Bağımlı Değişken Olması Durumunda Somer'in d Katsayısının En Yüksek Olabilirlik Tahmini
$\hat{d}_{XY}$	:	Y Bağımsız, X Bağımlı Değişken Olması Durumunda Somer'in d Katsayısının En Yüksek Olabilirlik Tahmini
f	:	Gözlenen Frekanslar
f'	:	Beklenen Frekanslar
f_i	:	i. Satırın Toplamı
f_j	:	j. Sütunun Toplamı
F	:	F Değeri
G	:	Goodman Kruskal'ın Gamma Katsayısı
$GAKT_{YX}$	:	Y Değişkeni Bağımlı Değişken Varsayıldığında Gruplar Arası Kareler Toplamı
$GAKT_{XY}$	:	X Değişkeni Bağımlı Değişken Varsayıldığında Gruplar Arası Kareler Toplamı
GKT_{YX}	:	Y Değişkeni Bağımlı Değişken Varsayıldığında Genel Kareler Toplamı
GKT_{XY}	:	X Değişkeni Bağımlı Değişken Varsayıldığında Genel Kareler Toplamı
GIKT	:	Guruplar İçi Kareler Toplamı
GASD	:	Gruplar Arası Serbestlik Derecesi
GISD	:	Gruplar İçi Serbestlik Derecesi
GAKO	:	Gruplar Arası Kareler Ortalaması
GIKO	:	Gruplar İçi Kareler Ortalaması
h	:	Normal Dağılım Eğrisi Altında Kalan Alanda p ve q'yu Ayıran Noktanın Ordinat Yüksekliği
i	:	i. Satır
j	:	j. Sütun
k	:	Satır ya da Sütun Sayısından Küçük Olanı (Karesel Olmayan Tablolor)
κ	:	Cohen'in Kappa Katsayısı
Max(R)	:	X Değişken Kategorileri İçerisinde En Yüksek Toplam Frekansı
Max(C)	:	Y Değişken Kategorileri İçerisinde En Yüksek Toplam Frekansı
n	:	Örneklem Genişliği
N	:	Toplam Gözlem Sayısı
n_i	:	i. Gruptaki Gözlem Sayısı
$n_{(max)i}$	:	i. Satırdaki En Büyük Değeri
$n_{(max)j}$	:	j. Sütundaki En Büyük Değeri

$\mathbf{n_{i+}^2}$	:	i. Satır Toplamının Karesi
$\mathbf{n_{+j}^2}$	:	j. Sütun Toplamının Karesi
P	:	Toplam Uyumlu Çift Sayısı
$\mathbf{P_a}$	:	Toplam Uyum Olasılığı
Pe(κ)	:	Şansa Bağlı Uyum Olasılığı
Q	:	Toplam Uyumsuz Çift Sayısı
p, q	:	Süreksiz Değişkendeki İki Kategorinin Oranları
P – Q	:	Uyumlu ve Uyumsuz Çiftlerin Farkı
r	:	Satır Sayısı
$\mathbf{r_b}$	:	Çift Serili Korelasyon Katsayısı
$\mathbf{r_{xy}}$	:	Pearson Korelasyon Katsayısı
$\mathbf{r_s}$	:	Spearman Rank İlişki Katsayısı
$\mathbf{r_{pq}}$	:	Nokta Çift Serili Korelasyon Katsayısı
$\mathbf{r_T}$	:	Kendall Tau Katsayısı
$\mathbf{R^2}$	:	Determinasyon (Belirlilik) Katsayısı
$\mathbf{R_{max}}$	:	En Büyük Toplam Satır Gözlem Sayısı
SD	:	Serbestlik Derecesi
$\mathbf{S_y}$	:	Sürekli Değişkenin Standart Sapması
t	:	t Değeri
$\mathbf{T_b}$	:	Kendall tau-b Katsayısı
$\mathbf{T_c}$	:	Kendall tau-c Katsayısı
$\mathbf{U_{(y/x)}}$	:	Y Bağımlı, X Bağımsız Değişken Olduğunda Hesaplanan Belirsizlik Katsayısı
$\mathbf{U_{(x/y)}}$	:	X Bağımlı, Y Bağımsız Değişken Olduğunda Hesaplanan Belirsizlik Katsayısı
$\mathbf{U_{(x,y)}}$	:	Simetrik belirsizlik katsayısı
$\mathbf{\bar{Y}_p, \bar{Y}_q}$	:	Süreksiz Değişkendeki İki Kategorinin Sürekli Değişkendeki Ölçümleri Ortalamaları
χ^2	:	Ki-kare
$\mathbf{x_{ij}}$	:	i. Muamele Grubundaki j. Tekerrüre Ait Gözlem Değeri
$\mathbf{\bar{x}_i}$	:	i. Grubun Ortalaması
$\mathbf{\bar{\bar{x}}}$	:	Ortalamaların Ortalaması
Z	:	Z Değeri
λ	:	Goodman Kruskal'ın Lambda Katsayısı
$\mathbf{\lambda_{X/Y}}$	:	Asimetrik Lambda için Y'nin X'e Göre Öngörü Hatası
$\mathbf{\lambda_{Y/X}}$	:	Asimetrik Lambda için X'in Y'ye göre Öngörü Hatası
Φ	:	Phi Katsayısı
$\mathbf{\eta_{YX}}$	:	Y Değişkeni Bağımlı Değişken Varsayıldığında Hesaplanın Eta Katsayısı
$\mathbf{\eta_{YX}}$	:	X Değişkeni Bağımlı Değişken Varsayıldığında Hesaplanan Eta Katsayısı
$\sum \mathbf{d_x^2}$	:	X Özelliklerine İlişkin Kareler Toplamı

$\sum \mathbf{d}_{\mathbf{y}}^2$	:	Y Özelliklerine İlişkin Kareler Toplamı
$\sum d_x d_y$	:	Çarpımlar Toplamı

1. GİRİŞ

Doğada bir değişkeni etkileyen birden fazla değişken olması sebebiyle araştırmalarda çoğunlukla birden fazla değişken/faktör ile çalışılmaktadır. Bu durum beraberinde değişkenler arasındaki ilişkilerin araştırılması gerekliliğini getirmektedir. Dolayısıyla değişkenler arasındaki etkileşimin yani ilişkinin da ortaya çıkarılması gerekmektedir. Değişkenler arasındaki ilişkilerin ortaya çıkarılmasında istatistik yöntem olarak en çok ilişki katsayılarından yararlanılır. Değişkenler arasındaki ilişkinin olup olmadığını varsa derecesinin ne olduğunu ve ilişkinin yönünü belirleyebilmek için ilişki katsayıları hesaplanmaktadır. Bu amaçla değişkenler arasındaki ilişkilerin ortaya çıkarılmasında kullanılabilecek çok sayıda ilişki katsayısı geliştirilmiştir.

Hesaplanması gereken ilişki katsayısını değişkenlerin tipini ve dağılım şeklini dikkate alarak belirlemek mümkündür. Kategorik tipteki değişkenler için kullanılacak ilişki katsayısı kategori sayısına ve kategoriler arasında sıralama olup olmamasına göre değişim göstermektedir. Örneğin; isimsel (nominal) değişkenlerde iki kategorilisubler için Phi katsayısı, ikiden fazla kategoriliiler için ise Cramer'in V katsayısı ve Lambda (λ) katsayısı kullanılmaktadır. Sıralı değişkenler arasındaki ilişkiler ise duruma göre Spearman rank, Kendall tau-b, Kendall tau-c veya Somer'in d katsayıları ile belirlenmektedir.

Bu ilişki katsayıları arasında hangi durumda hangisinin tercih edilmesi gerektiğine karar vermek araştırıcıların en büyük problemi haline gelmektedir. Pearson korelayon katsayısı uygulamada en yaygın kullanılan ilişki katsayısıdır. Ancak Pearson korelasyon katsayısı parametrik bir ilişki katsayısı olup kullanılabilmesi ve güvenilirliği bazı varsayımların sağlanmış olmasına bağlıdır. Bu varsayımlar;

- İki değişkeninde sürekli yapıda olması
- İki değişkeninde normal dağılımlı olması
- n≥10 olması
- Veri setinde sapan değer olmamasıdır.

Pearson korelasyon katsayısının varsayımlarının yerine gelmediği durumlarda uygulamada genel olarak Spearman rank ilişki katsayısı en çok önerilen ve kullanılan ilişki katsayısıdır. Aslında iki sıralı (ordinal) ölçekli kategorik değişken arasındaki ilişkiyi incelemek için geliştirilmiş olan Spearman rank ilişki katsayısının her deneme koşulundaki performansı yeterince bilinmemektedir. Normal olmayan veri gruplarında; Fowler (1987) yaptığı çalışmada Spearman rank ilişki katsayısının Pearson korelasyon katsayısına göre daha güçlü olduğunu bildirirken, Tuğran (2015) yaptığı çalışmada birçok deneme koşulunda Pearson, Winsorize ve Permütasyon tabanlı korelasyon katsayılarının, Spearman rank ve Kendall Tau korelasyon katsayılarına göre daha güçlü olduklarını bildirmiştir.

Bu tez çalışmasının iki temel amacı bulunmaktadır;

1. Kategorik değişkenler arasındaki ilişkilerin belirlenmesinde yaygın olarak kullanılan ilişki katsayıları hakkında detaylı bilgi vermek ve kendi içinde sınıflandırıp hangi durumlarda bu testlerin kullanılabileceğine dikkat çekerek veri tipine ve örneklem genişliğine göre uygun katsayıyı seçmektir.

2. Bazı korelasyon katsayılarının simülasyon yaklaşımı ile belirlenen deneme koşullarındaki performanslarını ampirik (deneye ve gözleme dayalı) olarak karşılaştırmaktır.

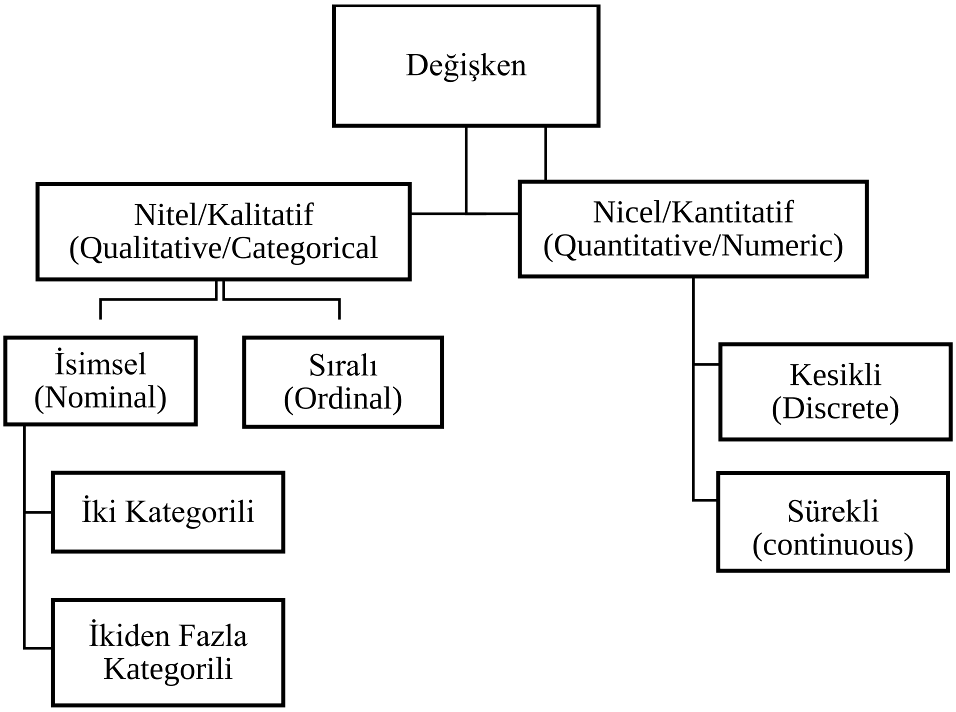
2. GENEL BİLGİLER

2.1 Değişken Kavramı ve Ölçme Seviyeleri

Değişken; deney ünitelerine veya bireylere ait olan sayılarak, ölçülerek, tartılarak analiz edilerek, veya isimsel olarak elde edilen ve birden çok değer alabilen bir özelliktir (Mendeş, 2012).

Değişkenler arası ilişkiler; bir özelliğin başka bir özelliği ya da özellikleri etkilemesi olarak tanımlandığı gibi aynı zamanda da bu özelliklerden etkilenmesidir. Bağımlı değişken, etkilenen ve bağımsız değişken ise etkileyendir (Sümbüloğlu ve ark., 1998).

Araştırıcının bağımlı değişken üzerinde etkisini gözlemlemek istediği değişken bağımsız değişkendir. Bağımlı değişken ise üzerinde bağımsız değişkenin etkisi test edilen değişkendir (Üstün, 2016).



Şekil 2.1 Değişkenlerin Veri Yapıları Açısından Sınıflandırılması (Kaşko, 2007)

Değişkenler kategorik (Nitel) ve sayısal (nicel) değişkenler olarak ayrılır. (Şekil 2.1). Sayısal değişkenler elde ediliş şekillerine göre, sürekli ve kesikli değişkenler olmak üzere ikiye ayrılır. Sürekli değişkenlerde; gözlem değerleri ölçerek, tartarak ya da analiz ederek elde edilir. Kesikli değişkenlerde ise saymak üzere elde edilmektedir. Sürekli değişkenlere ağırlık, boy, yaş, sütteki % protein miktarı örnek gösterilebilir. Kesikli değişkenlere ise ailedeki çocuk sayısı, ağızdaki diş sayısı ve böcekteki segment sayısı örnek olarak verilebilir (Kaşko, 2007).

Kategorik (nitel) değişkenler, isimsel (nominal) ya da sıralı (ordinal) veri özelliği gösteren değişkenlerdir. İsimsel değişkenler iki şıklı ve ikiden fazla şıklı olup, belirlenmiş olan şıklar arasında küçüklük veya büyüklük sırası söz konusu değildir. Sıralı değişkenler için yukarıdan aşağıya ya da aşağıdan yukarıya doğru küçüklük ve büyüklük sırası vardır. İki şıklı değişken; kadın-erkek, evet-hayır, bozuk-sağlam, var-yok gibi yalnızca iki değer alabilen değişkendir. Çok şıklı değişken ise ikiden fazla şıkkı bulunan değişkendir. Sayısal değişkenler, kesikli ve sürekli değişkenler olarak ayrılır. Aralık (interval) ve oransal (ratio) ölçek verilerinde sonsuz sayıda değer alabilen değişkenler sürekli değişkenlerdir. Sürekli değişkenlere gruplandırılmamış yaş değişkeni, ücret değişkeni, üretim hacmi değişkeni, boy ve kilo değişkeni gibi örnekler verilebilir. Araştırmacının veri toplarken sürekli değişken veri biçimini tercih etmesi, bu verilerin kategorik değişken biçimine dönüştürülmesinde kolaylık sağlar. Sürekli değişkenler olan aralık ve oransal değişkenlere istatistik analiz tekniklerini uygulamak mümkündür (Şencan, 2007).

Kategorik değişkenleri, yalnızca sınırlı sayıda değer veya kategori kullanılarak ölçülebilen değişkenler olarak tanımlarız. Bu tanım, kategorik değişkenleri, sonsuz sayıda değer alabilen sürekli değişkenlerden ayırır (Powers ve Xie, 2008).

Ölçme seviyeleri duyarlılıklarına göre dört gruba ayrılır (Gamgam ve Altunkaynak, 1989).

1. İsimsel (nominal) ölçme seviyesi
2. Sıralama (ordinal) ölçme seviyesi
3. Aralık (interval) ölçme seviyesi
4. Oransal (ratio) ölçme seviyesi

2.1.1 İsimsel (Nominal) Ölçme Seviyesi

İsimsel ölçme seviyesinde birimlere ad verilerek değerlendirme yapmayı sağlayan bir ölçme seviyesidir. Birimlere sayı, simge veya isimler verilebilir. Tercih edilen marka, cinsiyet, göz rengi, meslek grupları, tedavide kullanılan ilaç türü, tarımda kullanılan gübre çeşidi isimsel ölçme seviyesinde ölçülebilen değişkenlere örnek olarak verilebilir (Gamgam ve Altunkaynak, 1989).

İsimsel ölçme seviyesinde birimlere verilen sayılar sadece isim olarak düşünüldüğünden grupları birbirinden ayırmak amacıyla kullanılır. Bu nedenle verilen sayılar bir önem sıralaması ya da büyüklük ve küçüklük sırası değildir.

2.1.2 Sıralama (Ordinal) Ölçme Seviyesi

İsimsel ölçme seviyesinden daha hassas olan sıralama ölçme seviyesidir. Sıralama ölçme seviyesinde birimlere verilen sayılar birimlerin birbirlerinden farklı olduklarını gösterir. Aynı zamanda ilgili değişken bakımından birimlerin büyüklük sırasını ya da daha önemli olduğunu belirlemek amacıyla da kullanılır. Firma büyüklükleri, ekonomik durum, televizyon programlarını izleme sıklığı, sosyal statü, başarı durumu gibi değişkenler sıralama seviyesinde ölçülebilecek değişkenler için örnek olarak düşünülebilir (Gamgam ve Altunkaynak, 1989).

Sıralı değişkenlerde ölçüm seviyeleri arasında mesafeleri belirli olmayan bir sıralama vardır. Bu nedenle önem ve büyüklüklerine göre sıralanırken miktarını yani kategoriler arasında ne kadar bir fark olduğu bilgisinin belirtilmediği ölçme seviyesidir. Üzerinde durulan özelliğin önem ya da büyüklüklerine göre küçükten büyüğe ya da büyükten küçüğe doğru sıralama yapılabilir. Sıralı değişkenler için başarı durumunu; kötü, orta, iyi, çok iyi ve sınavın zorluk derecesi; çok zor, zor, kolay, çok kolay şeklinde örnek olarak gösterilebilir (Kaşko, 2007).

2.1.3 Aralık (İnterval) Ölçme Seviyesi

Aralık ölçme seviyesi, isimsel ve sıralama ölçme seviyelerine göre daha hassas bir ölçme sağlar. Hava sıcaklığı, yalnızlık duygusu, zeka derecesi gibi değişkenler aralık seviyesinde ölçülebilen değişkenlere örnek olarak verilebilir (Gamgam ve Altunkaynak, 1989).

Bu ölçme seviyesinde verilerle çarpma, bölme, çıkarma, toplama işlemleri yapılabilir. Ancak isimsel ve sınıflandırma ölçme seviyelerinde bu matematiksel işlemler yapılamaz.

Aralık ölçme seviyesinde, başlangıç noktasının seçimi keyfi olarak belirlenir. Belirlenen bir başlangıç noktası ile yani sabit olmayan noktadan başlayıp eşit aralıklara ayrılarak sıralanır (Taştan, 2013).

Aralık ölçek için sıfır noktası görecelidir. Örneğin gün içindeki sıcaklığın 0 derece olduğu bilinmesi, -3 dereceden daha fazla ısı olduğunu ifade eder. Sıcaklığın 0 derece olması hiç olmadığını belirtmez (Karagöz, 2010b).

2.1.4 Oransal (Ratio) Ölçme Seviyesi

Sıfır noktasının ya da standart bir başlangıç noktasının olmaması aralık ölçme seviyesinin, temel özelliği olduğu gibi aynı zamanda bu eksikliğidir. Oransal ölçme seviyesinde sıfır noktası mutlaktır (Gamgam ve Altunkaynak, 1989).

Oransal ölçme seviyesi ile ölçülen özelliğin olmadığını sıfır sayısı belirtir. Bu ölçme seviyesinde matematiksel tüm işlemler yapılabilir.

2.2 Korelasyon (İlişki) Kavramı

Korelasyon; iki tesadüfi değişken arasındaki doğrusal bir ilişkinin ölçüsüdür. Bu ilişkiyi ölçen katsayı korelasyon katsayısıdır (Demür, 2014).

Değişkenler arasındaki ilişkinin yönü “pozitif”, “negatif” ve “ilişki yok” şeklinde tanımlanır. Pozitif ilişki, bir değişkenin artmasıyla birlikte diğer değişkende de artma meydana gelmesi veya azaldığında diğerinin de azalmasıyla birlikte aynı yönde olan ilişki türüdür. Negatif ilişki, bir değişken arttığında diğer değişkeninde ters yönde azalması veya azaldığında diğer değişkeninde artması olarak tanımlanır. Bir değişkenin artması ya da azalmasıyla, diğer değişkende de hem azalmanın hem de artmanın görülmesi ilişkinin olmadığı yani yokluğu şeklinde tanımlanır (Sümbüloğlu ve ark., 1998).

İlişki olduğu zaman, saçılma diyagramı bir daireden veya yatay bir elipsten farklı bir şekil alır ve regresyon doğrusu yatay bir doğru özelliği göstermez. Bir saçılma diyagramının şekli ve yönelimi, veriler tarafından şekillendirilmiş ilişkinin özelliğini yansıtır. Bu özellikleri özetlemek için, korelasyon katsayısı olarak bilinen istatistik hesaplanır. Korelasyon katsayısı, iki değişken arasındaki ilişkinin önemli özelliklerini açıklayan ve özetleyen bir sayıdır. Bir ilişkinin iki önemli özelliği ilişkinin tipi ve ilişkinin gücüdür. İlişki tipi bir veri grubunda değişkenlerden birinin değişimiyle diğerinin değişiminde gözlemlenen yön tarafından belirlenir.

Çizelge 2.1’de gösterildiği gibi doğrusal ve doğrusal olmayan iki ilişki tipi vardır. Doğrusal bir ilişkiyi, bir doğru çizgi olan regresyon doğrusu özetler. Bir daire ya da yatay bir elips oluşmayıp, bir doğru ile gösterilebildiğinde doğrusal ilişkinin varlığından bahsedilebilir. Saçılma diyagramları pozitif ve negatif olarak iki tür doğrusal ilişkiyi gösterir. Pozitif doğrusal ilişkide değişkenlerden biri artarken diğerinde de bir artma veya biri azalırken diğerinde de azalma olduğu görülür. Negatif doğrusal ilişki de ise değişkenlerden birinin artmasıyla diğerinde bir azalma veya biri azalırken diğer değişkende artma meydana gelir. Doğrusal olmayan ilişkiye "nonlinear" veya "curvilinear" adı da verilir. Veriler bir doğru ile özetlenemez. Yani doğrusal olmayan ilişkide değişkenlerden birinde olan değişim ile birlikte diğer değişkende sadece artma veya sadece azalma olmaz. Örneğin bir değişkenin değişimiyle birlikte diğer değişkende artma gösterirken belli bir noktadan sonra azalma görülebilir. Böylece hem artma hem de azalmanın olabileceği doğrusal olmayan bir ilişkinin varlığı söz konusudur (Büyüköztürk ve ark., 2000).

Çizelge 2.1 Korelasyon Katsayıları (Anonim, 2015)

KORELASYON			
Doğrusal olan ilişki			**Doğrusal olmayan ilişki**
Paremetrik olmayan		**Paremetrik olan**	Eta
Sıralama (Ordinal)	**İsimsel (Nominal)**	**Sürekli (Ölçme)**	
Gamma	Ki-kare	Pearson	
Kendall tau-b	Phi		
Kendall tau-c	Cramer'in V		
Spearman Rank	Lambda		
Somer'in d	Contingency (Bağımlılık)		
	Belirsizlik (Uncertainty)		

3. KATEGORİK DEĞİŞKENLER ARASI İLİŞKİ KATSAYILARI

Veriler sürekli olduğunda yani aralık, oransal veya kategorik ise isimsel, sıralama olup olmadığına, istatistiksel analiz yapılmadan önce dikkat edilmelidir. Ayrıca bunlardan en güçlüsünün oransal ölçek olduğu görülür. Sonra bunu aralık, sıralama ve isimsel ölçekleri takip eder. Oransal ölçeğinden aralık ölçeğine, aralık ölçeğinden sıralama ölçeğine, sıralama ölçeğinden isimsel ölçeğe geçiş yapılabilir. Böyle bir geçişte her birisi için bilgi kaybı olabileceğinden isimselden sıralamaya, sıralamadan aralık ölçeğe, aralık ölçekten de oransal ölçeğe doğru dönüşüm mümkün değildir. Çünkü istatistiksel araştırmalar için toplanan veriler kaydedilirken daha güçlü bir ölçekten faydalanılmak istenir. Parametrik testler sürekli verilerde, parametrik olmayan testler de kategorik verilerde kullanılması uygundur. Parametrik testlerin uygulanabilmesi için bazı varsayımların sağlanması gereklidir. Bu varsayımlar; verilerin normal dağılım göstermesi, varyansların homojen olması ayrıca her teste özgü farklı koşullarında sağlanması aranır. Parametrik olmayan testler ile kıyaslandığında parametrik testlerin daha güçlü olduğu görülür. Yani parametrik olmayan testlerin daha az güçlü olmasının sebebi II. tip bir hata olasılığının parametrik olmayan testte daha büyük olmasıdır. Normal dağılım göstermeyen ve sıralama verilerde parametrik test uygulanması sakıncalı varsayılırken, isimsel dağılım gösteren veriler için parametrik olmayan testler uygulanması durumunda pek hatalı olduğu görülmez. İstatistiksel testlerin uygulanması için öngörülen her testin kendine ait koşulların neler olduğu ve bu koşullara bağlı olarak verilerin uygunluğunun nasıl belirlenmesi gerektiği iyi bilinmelidir. Koşulların sağlandığına dair net bir bilgi olmaması durumunda verilerin analizi için parametrik olmayan testlerin kullanılması, doğru sonuçlar elde edebilmek açısından önemlidir. Parametrik testlerin yapılabilmesi için gerekli görülen varsayımların sağlanmasına rağmen parametrik olmayan testlerin kullanılması, parametrik testlerin avantajlarının göz ardı edilmiş olmasına sebep olur. Kategorik ve sıralama ölçekli veriler için en uygun olan ve parametrik testlerin varsayımları sağlanmadığında başvurulan testler parametrik olmayanlardır. Çünkü bunlar daha az koşulların varlığında uygulanabilir (Kartal, 2006; Kalaycı ve ark., 2010; Karagöz, 2010c).

Değişkenler arasındaki ilişkilerin belirlenmesinde kullanılacak ilişki katsayısı veri tipine ve kullanılacak katsayının varsayımlarına göre değişim göstermektedir. Örneğin; sürekli tipteki değişkenler için varsayımları karşılaması durumunda Pearson korelasyon katsayısı, değişkenler arasındaki ilişkilerin ortaya çıkarılması için en uygun olanıdır. Kategorik tipteki değişkenler için ise kullanılacak ilişki katsayısı kategori sayısına ve kategoriler arasında sıralama olup olmamasına göre değişim göstermektedir.

İlişki katsayılarının sınıflandırılmasında, değişken tipleri ve ilişkili olan değişken sayısı önemlidir. İki (bivariate) ve çoklu (multiple) değişken olarak ilişki katsayıları vardır. İki değişken arasındaki ilişkiyi, iki değişkenli ilişki katsayıları belirlerken bir değişken ile değişkenler seti arasındaki ilişkileri ise çok değişkenli ilişki katsayıları belirler. İlişki yönü hakkında, iki değişkenli ilişki katsayılarının bazıları bilgi vermediğinde simetrik ilişkiler olarak tanımlanır. İlişkinin yönü hakkında bilgi veren ilişki katsayıları ise simetrik olmayan ilişkilerdir (Öztuna ve ark., 2008; Aslan ve ark., 2019).

Bağımlı ve bağımsız değişken ayrımı yapılmadığında hesaplanan katsayılar simetrik; bağımlı ve bağımsız değişken ayrımı yapıldığında hesaplanan katsayılar simetrik olmayan ilişki katsayılarıdır (Siegel ve Castellan, 1988; Öztuna ve ark., 2008).

X değişkenindeki 1 birimlik artış ve 1 birimlik azalışın, Y değişkeni üzerindeki etkisi aynı olduğunda ilişki simekriktir. Yani bağımlı ve bağımsız değişken ayrımı yapılmayan durumdur. Bu farklılaştığında artık simetrik olmayan ilişkidir. Simetrik olmayan ilişkide değişkenlerden biri, bağımlı değişken olarak tanımlanır (Köse, 2020).

Çizelge 3.1’de gösterildiği gibi simetrik olan isimsel değişkenler için kullanılan ilişki katsayıları Phi, Contingency, Cramer’in V iken sıralı değişkenler için Gamma, Kendall tau-a, Kendall tau-b, Kendall tau-c, Spearman rank katsayılarıdır. Simetrik olmayan isimsel değişkenler için kullanılan Lambda, Belirsizlik (Uncertainty) ve sıralı değişkenler için ise Somer’in d katsayılarıdır. Lambda, Belirsizlik ve Somer’in d ilişki katsayıları değişkenlerden birini bağımlı diğerini de bağımsız değişken olarak belirleyebildiğinden simetrik olmayan katsayılardır.

Çizelge 3.1 Simetrik ve Simetrik Olmayan İlişki Katsayıları (Lebe, 2006; Öztuna ve ark., 2008)

Değişkenler	**Simetrik ilişki katsayıları**	**Simetrik olmayan ilişki katsayıları**
İsimsel (Nominal) değişkenlerde	Phi	Lambda
	Contingency	Belirsizlik (Uncertainty)
	Cramer’in V	
Sıralı (Ordinal) değişkenlerde	Gamma	Somer’in d
	Kendall tau-a	
	Kendall tau-b	
	Kendall tau-c	
	Spearman rank	

Ki-kare testi iki kategorik değişken arasındaki ilişkinin test edilmesinde kullanılabilir. Ancak değişkenler arasında ilişki olup olmadığını belirlerken yönünü ve büyüklüğünü açıklayamaz. Bunu ilişki katsayılarıyla belirlemek mümkündür. Bir kategoriye ait beklenen değer küçük bir sayı olduğunda Ki-kare kullanımı hatalı sonuç verebilir. Yani beklenen değer küçüldükçe Ki-karenin değeri daha çok büyür. Böylece sıfır hipotezi reddedilir. Bu sebeplere bağlı olarak değişkenler arasında ilişkinin olup olmadığı aranırken, yönünü ve büyüklüğünün de belirlenebilmesi açısından Ki-kare testinin yanı sıra diğer ilişki katsayılarının kullanılması avantajlı olacaktır (Daniel, 1990; Oktay, 2003; Altunışık ve ark., 2006; Kartal, 2006; Karagöz, 2010a).

Beklenen değerin küçülmesine bağlı olarak Ki-kare değerinin büyümesi sonucu sıfır hipotezi reddedilir. Bunun için Ki-kare testi yaparken, beklenen değerlerin 1’den küçük olmaması ve %20’den fazlasının da beşten küçük olmaması gerekir. Ki-kare bağımsızlık testi örnek genişliğinden de bağımsız değildir. Ki-kare kökenli olan Phi, Contingency ve Cramer’in V ilişki katsayıları ile örnek genişliğinin, test istatistiği üzerindeki etkisi yok edilir. Ki-kare kökenli bu ilişki katsayıları, Ki-kare bağımsızlık test sonucu önemli bulunduğunda kullanılır. Ki-kare kökenli ilişki katsayıları kontenjans tablosunun boyutlarına, satır ve sütun toplamlarına karşı hassas ölçüler olduklarından yorumlanmalarında ciddi güçlükler bulunur. Ki-kare kökenli ilişki katsayılarının bu tür olumsuzluklarını gidermek için tahmin hatasını azaltmaya dayalı Goodman Kruskal'ın Lambda ve Theil'in belirsizlik (Uncertainity) katsayısı gibi ilişki katsayıları geliştirilmiştir (Karagöz, 2010b).

Çizelge 3.2 Değişken Türlerine Göre İlişki Katsayıları (Borg ve Gall, 1989; Balcı, 1995; Karagöz, 2010a; Aslan ve ark., 2019)

İlişki Katsayısı	**Birinci Değişken**	**İkinci Değişken**	**Açıklama**
Sayısal değişkenler için ilişki katsayıları			
Pearson	Sürekli	Sürekli	Değişkenlerin her ikisi de normal dağılım gösterdiğinde, nokta grafiği linear ve/veya eliptik bir yapıda olduğunda kullanılır.
Spearman rank	Sürekli Sürekli Kesikli Sıralı (Ordinal)	Sürekli Kesikli Kesikli Sıralı (Ordinal)	Değişkenlerden en az biri normal dağılım gösterdiğinde, nokta grafiği linear ve/veya eliptik bir yapıdan ayrılışlar gösterdiğinde ve/veya gözlem sayısı az olduğunda kullanılır. Ayrıca her iki değişkenin gerçek sıralı değişken olması durumunda kullanılır.
İsimsel (Nominal) değişkenler için ilişki katsayılar			
Phi	İsimsel (Nominal)	İsimsel (Nominal)	İki değişken için de iki kategorili olup, veriler 2×2 kontenjans tablo ile özetlendiğinde Phi katsayısı kullanılır.
Cramer'in V	İsimsel (Nominal)	İsimsel (Nominal)	Satır ve sütun sayısı eşit olan R×C kontenjans tablolarda kullanılabilir. 2×2'lik tablolardan büyük olmalıdır. Cramer'in V katsayısı tablo büyüklüğünden etkilenmez. Ki-kare kökenli ilişki katsayısıdır.
Contingency (Bağımlılık)	İki ya da Daha Çok Kategori	İki ya da Daha Çok Kategori	Ki-kare kökenli katsayıdır.
Lambda	İsimsel (Nominal)	İsimsel (Nominal)	2×2'lik tablolarda ve daha büyük tablolarda kullanılabilir.
İsimsel (Nominal) bir değişken ile diğeri kesikli/sürekli olan bir değişken için ilişki katsayıları			
Nokta Çift Serili Korelasoyon Katsayısı	İkili İsimsel (Binary Nominal)	Sürekli Kesikli	Bir değişkenin iki ya da çok kategorili isimsel, diğer değişkenin sayısal veri türünde olduğunda kullanılır.
Çift Serili Korelasoyon Katsayısı	İkili İsimsel (Binary Nominal)	Sürekli Kesikli	Bir değişkenin iki kategorili isimsel, diğer değişkenin sayısal veri türünde olduğu durumlarda kullanılır.
Sıralı (Ordinal) değişkenler için ilişki katsayıları			
Goodman ve Kruskal'ın Gamma	Sıralı (Ordinal)	Sıralı (Ordinal)	Bu katsayı değişkenlerin sıralı ve kontenjans tablo şeklinde olması durumlarında kullanılır.
Kendall Tau	Sıralı (Ordinal)	Sıralı (Ordinal)	Karesel kontenjans tablolar için kullanılabilen Kendall Tau katsayısı, 10'un altındaki sıralamalarda spearman rank yerine tercih edilir.
İsimsel (Nominal) ve Aralık (İnterval) ölçekli ilişki katsayıları			
Eta	İsimsel (Nominal)	Aralık (İnterval)	Doğrusal olmayan ilişkileri tespit için kullanılır. Bağımsız değişkende isimsel, bağımlıda ise en az aralık ölçeği ile ölçülmüş olması koşulu aranır.

3.1 İsimsel (Nominal) Değişkenler Arasındaki İlişkinin Belirlenmesinde Kullanılan İlişki Katsayıları

3.1.1 Phi Katsayısı

Parametrik olmayan testlerden olan Phi katsayısı, 2×2'lik tablolarda verilerin korelasyon katsayısını bulmak için kullanılan isimsel ölçekle elde edilen bir testtir (Terzi, 2019).

İsimsel iki sonuçlu ve iki değişken arasındaki ilişkinin büyüklüğünü ölçen bir ilişki katsayısıdır. Phi katsayısı, Ki-kare istatistiğinin önemli olduğu durumda kullanılmalıdır. Örneklem genişliğinin az olduğu durumlarda Phi katsayısının önemlilik testi Fisher'in Ki-kare testi ile yapılmalıdır. Değişkenler arasında aynı yönlü ilişki varsa verilerin çoğunluğu köşegen hücrelerinde yer alırken, ters yönlü ilişki durumu söz konusu olduğunda köşegen dışı hücrelerde yer alır. Değişkenler arasında ilişki olmadığında yani birbirlerinden bağımsızlarsa ilişki katsayısı 0'a eşit olur. İlişki katsayısı +1 ya da -1 ise değişkenler arasında ilişki olduğu söylenebilir (Siegel ve Castellan, 1988; Öztuna ve ark., 2008).

İlişki katsayısı +1 ise değişkenler arasında pozitif ilişki, -1 ise değişkenler arasında tam negatif ilişki vardır. Boyutları 2×2'lik tablolardan büyükse Phi katsayısı çoğunlukla üst sınırı yakın değerleri yakalayamamaktadır. 2×2 boyutundaki tablolardan elde edilen Phi katsayısı, Cramer'in V katsayısı ile aynı değeri aldığı görülür (Koçyiğit, 2016).

İki değişkende ikili olduğunda değişkenler arasındaki ilişki Phi (φ) katsayısından yararlanılarak belirlenir. Buna göre Phi katsayısı aşağıdaki gibi hesaplanır (Mendeş, 2012).

Çizelge 3.3 Çapraz Tablo

X değişkeni	Y değişkeni		Toplam
	0	1	
1	A	B	A+B
0	C	D	C+D
Toplam	A+C	B+D	N

$$\varphi = \frac{(AD - BC)}{\sqrt{(A+B)(C+D)(A+C)(B+D)}} \tag{2.1}$$

Bu şekilde de hesaplanan Phi katsayısı -1 ile +1 arasındaki değerleri alarak değişkenler arasındaki doğrusal ilişkinin yanında yönü hakkında da bilgi verir.

$$f' = \frac{f_i f_j}{N} \tag{2.2}$$

$$\chi^2 = \sum_{k=1}^{n} \frac{(f - f')^2}{f'} \tag{2.3}$$

$$\varphi = \sqrt{\frac{\chi^2}{N}} \tag{2.4}$$

Şeklinde olup burada,

f' : Beklenen frekanslar,
f : Gözlenen frekanslar,
f_i : i. satırın toplamı,
f_j : j. sütunun toplamı,
N : Toplam gözlem sayısı,
χ^2 : Ki-kare test istatistiğidir.

Phi katsayısı aynı zamanda bu şekilde de hesaplanabilir. Fakat bu yöntemle hesaplandığında Phi katsayısı sadece 0 ile 1 arasındaki değerleri alabilir.

Phi katsayısı, Contingency ve Cramer'in V katsayıları Ki-kare kökenli katsayılar olduğu için önemlilik testi Ki-kare değeri hesaplanarak yapılabilir. Hesaplanan Ki-kare değeri, R satır ve C sütun sayısı olmak üzere (R-1)(C-1) şeklinde hesaplanan serbestlik dereceli Ki-kare dağılımını gösterir. Hesaplanan Ki-kare değeri, serbestlik derecesi dikkate alınarak bulunan Ki-kare tablo değerinden daha büyük olduğunda sıfır hipotezi reddedilir. Böylece değişkenler arasında önemli bir ilişkinin ya da bağımlılığın olduğuna karar verilir. Bunun sonucu olarak da bu ilişkinin derecesi Ki-kare kökenli olan Phi, Contingency ve Cramer'in V katsayıları ile bulunur. Ki-kare değeri eğer Ki-kare tablo değerinden küçük ise sıfır hipotezi kabul edilerek önemli bir ilişkinin ya da bağımlılığın olmadığına karar verilir (Mendeş, 2012).

3.1.2 Cramer'in V Katsayısı

Cramer'in V katsayısı, isimsel iki değişken arasındaki ilişkinin gücünü yorumlayabilmemize olanak sağlar. Ki-kare testi anlamlı ise Cramer'in V katsayısının da anlamlı olduğu belirlenir. Ki-kare testi anlamlı olmadığında da aynı şekilde Cramer'in V katsayısı da anlamlı değildir. İki değişken arasında ilişki olmadığı zaman Cramer'in V katsayısı 0'a eşittir (Öztuna ve ark., 2008).

Tabloda satır ve Sütun sayılarının eşit (R=C), olması şartıyla Cramer'in V testi uygulanabilir. Yani tabloların 3×5, 2×3, 4×3 yerine 3×3, 4×4 şeklinde olması gerekir (Terzi, 2019).

Tablodaki satır ve sütun sayısı eşit ve katsayı değeri 1 olduğunda değişkenler arasında çok güçlü bir ilişkinin varlığıyla yorumlanır. Fakat satır ve sütun sayısı eşit olmadığında katsayı 1'e eşit olsa bile değişkenler arasında sadece bir yönde güçlü ilişkinin var olduğu görülür. Yani satır sayısı sütundan fazla (R>C) olduğunda, değişkenler arasında tersine bir ilişki vardır (Siegel ve Castellan, 1988; Alpar, 2001; Öztuna ve ark., 2008).

Cramer'in V testinin özetle uygulanabilmesi için; satır ve sütun sayısı eşit olan 2×2'lik tablolardan büyük ve çapraz tablolar olmalı, veriler isimsel ölçekle ölçülmüş olmalı, örnekler popülasyondan tesadüfi seçilmiş olmalıdır (Levin ve Fox, 1991; Terzi, 2019).

Cramer'in V katsayısı aşağıdaki gibi hesaplanır (Hamarat, 2017).

$$V = \sqrt{\frac{\chi^2}{N(k-1)}} \tag{2.5}$$

Şeklinde olup burada,

N : Toplam gözlem sayısı,
k : Karesel olmayan tablo için satır ya da sütun sayısından küçük olanı,
χ^2 : Ki-karedir.

Cramer'in V katsayısı önem testi Ki-kare istatistiği ile aşağıdaki gibi yapılır.

$$f' = \frac{f_i f_j}{N} \tag{2.6}$$

$$\chi^2 = \sum_{k=1}^{n} \frac{(f - f')^2}{f'} \tag{2.7}$$

Hesaplanan Ki-kare değeri, R satır ve C sütun sayısı olmak üzere (R-1)(C-1) şeklinde hesaplanan serbestlik dereceli Ki-kare dağılımını gösterir. Hesaplanan Ki-kare değeri, serbestlik derecesi dikkate alınarak bulunan Ki-kare tablo değerinden daha büyük olduğunda sıfır hipotezi reddedilir (Mendeş, 2012).

Ki-kare hesaplanan değer, Ki-kare tablo değerinden küçük bulunursa H_0 hipotezi kabul edilerek değişkenler arasında bir ilişkinin olmadığına karar verilebilir. İki değişken arasında ilişkinin bulunma veya bağımsızlığın olma durumu bu şekilde kontrol edilebilir (Üçkardeş, 2006).

3.1.3 Goodman Kruskal'ın Lambda Katsayısı

Bağımlı ve bağımsız değişkenlerin birbirlerini etkiledikleri hata oranları üzerinde hesaplama yapar. Birkaç kategoriyi isimsel ölçekle karşılaştıran bir ilişki katsayısıdır. Lambda 0 ile 1 arasında değer alır. Verilerin isimsel ölçekle alınmış olması, 2×2 ve daha büyük olan çapraz tabloların kullanılması ve örneklerin tesadüfi olarak seçilmesi varsayımlarıyla Lambda (λ) testi kullanılabilir (Levin ve Fox, 1991; Üçkardeş, 2016).

Lambda katsayısı, 0 olduğu durumda bağımsız değişkenden faydalanarak bağımlı değişkeni önceden tahmin edemeyeceğimiz anlaşılır. Ancak lambda katsayısı 1 olursa bağımsız değişkenin, bağımlı değişken için her bir kategorisini çok iyi belirlediğini ve tahminin doğru yapıldığını gösterir. Ayrıca Lambda katsayısı 0.5 olursa bağımsız değişken dikkate alınarak yapılan tahminlerin anlamlı olduğu görülebilir (Terzi, 2019).

Ki-kare değerinden bağımsız olan lambda katsayısı hatadaki orantılı azalma fikrine dayalıdır. Lambda katsayısı bir bağımsız değişken bilindiğinde, diğer bağımlı değişkenin tahminindeki hatanın göreceli azalmasının ölçütüdür. Bir Ki-kare tablosundan, üç farklı lambda değeri hesaplanabilir. Birincisinde sütun değişkeni tahmin edici olarak kullanılır. İkincisin de ise satır değişkeni tahmin edici olarak kullanılır (Öztuna ve ark., 2008).

Lambda katsayısı, Gutman'ın tahmin katsayısı olarak da adlandırılır. Lambda katsayısı; Phi ve Cramer'in V katsayılarına benzer olarak birkaç grup ya da kategorinin isimsel ölçekle karşılaştıran bir ilişki katsayısıdır. Lambda katsayısının bu testlerden farkı asimetrik yapıya sahiptir. Buna karşılık istenirse simetrik olarak da hesaplama yapmaya olanak sağlar (Levin ve Fox, 1991; Terzi, 2019).

Goodman Kruskal'ın Lambda Katsayısı aşağıdaki gibi hesaplanır (Üçkardeş, 2016).

Asimetrik Lambda için Y'nin X'e göre öngörü hatası;

$$\lambda_{X/Y} = \frac{\sum_{J=1}^{C} n_{(max)j} - Max(R)}{N - Max(R)} \tag{2.8}$$

Asimetrik Lambda için X'in Y'ye göre öngörü hatası;

$$\lambda_{Y/X} = \frac{\sum_{i=1}^{R} n_{(max)i} - Max(C)}{N - Max(C)} \tag{2.9}$$

Simetrik Lambda için;

$$\lambda = \frac{\sum_{i=1}^{R} n_{(max)i} + \sum_{j=1}^{C} n_{(max)j} - R_{max} - C_{max}}{2N - R_{max} - C_{max}} \tag{2.10}$$

Şeklinde olup burada,

λ	:	Goodman Kruskal'ın Lambda katsayısı,
$n_{(max)j}$	:	j. sütundaki en büyük değeri,
$n_{(max)i}$	:	i. satırdaki en büyük değeri,
N	:	Toplam gözlem sayısı,
Max(R)	:	X değişken kategorileri içerisinde en yüksek toplam frekansı,
Max(C)	:	Y değişken kategorileri içerisinde en yüksek toplam frekansı,
R_{max}	:	En büyük toplam satır gözlem sayısı,
C_{max}	:	En büyük toplam sütun gözlem sayısıdır.

3.1.4 Pearson'un Kontenjans (Contingency) Katsayısı

Örneklem büyüklüğünden bağımsız bir ilişki ölçüsü olan Contingency katsayısı 0 ile 1 arasında değerler alır. Değişkenler arasında ilişkinin bulunmaması durumunda katsayı değeri 0 olur. Mükemmel bir ilişkinin varlığında ise bu katsayı 1 değerini alır (Anonim, 2019a).

Araştırmacılardan bazıları 5×5 boyutundan küçük olan çapraz tablolarından elde edilen Contingency katsayılarının güvenilir olmadığını ve bu sebepten kullanılmaması gerektiğini belirtir (Oktay, 2003; Arıcıgil Çilan, 2009; Karagöz, 2010b; Blaikie, 2013; Nakip, 2013).

Contingency Katsayısı (2.11) numaralı formüldeki gibi hesaplanırken sadece iki özelliğin arasındaki bağımlılığın derecesini verir ve yönünü belirtmez. İki özellik bağımlı olsa bile katsayı 1 olamayacağından R×C tablolarında alabileceği en büyük değere göre contingency katsayısında düzeltme yapılmalıdır. İki yanlı tablolarda satır ve sütunların yer değiştirmesinden dolayı CC_{max1} ve CC_{max2}'nin ortalaması olarak CC_{max} (2.14) numaralı formüldeki gibi hesaplanır. Katsayının alabildiği en büyük değerinin hesaplanmasıyla birlikte son olarak düzeltilmiş Contingency katsayısı (2.15), (2.16), (2,17) numaralı eşitliklerden uygun olanıyla aşağıdaki gibi bulunur (Kocabaş ve ark., 2013).

$$CC = \sqrt{\frac{\chi^2}{N+\chi^2}} \tag{2.11}$$

$$CC_{max1} = \sqrt{\frac{r-1}{r}} \qquad \text{R<C ise} \tag{2.12}$$

$$CC_{max2} = \sqrt{\frac{c-1}{c}} \qquad \text{R>C ise} \tag{2.13}$$

$$CC_{max} = \frac{CC_{max1} + CC_{max2}}{2} \tag{2.14}$$

$$CC_d = \frac{CC}{CC_{max1}} \tag{2.15}$$

$$CC_d = \frac{CC}{CC_{max2}} \tag{2.16}$$

$$CC_d = \frac{CC}{CC_{max}} \tag{2.17}$$

3.1.5 Theil'in Belirsizlik (Uncertainity) Katsayısı

Simetrik olmayan bir katsayı olan Theil'in Belirsizlik katsayısı 0 ile 1 arasında değişen değerler alır. Belirsizlik katsayısı 0 olduğunda değişkenler arasında bağımsızlık vardır. Bağımlı değişkendeki değişmeyi, bağımsız değişken tam olarak açıklayabildiğinde bu katsayı 1 değerini alır. Bağımlı ve bağımsız değişkenlerin fonksiyonları birbirleriyle yer değiştirdiğinde farklı bir katsayı elde edilir (Astola ve Virtanen, 1981; Lebe, 2006).

$U_{y/x}$ katsayısı; Y değişkeni bağımlı, X değişkeni bağımsız değişken olarak varsayıldığında (2.21) numaralı formül yardımıyla hesaplanır. Benzer şekilde $U_{x/y}$ katsayıs; X değişkeni bağımlı, Y değişkeni bağımsız değişken varsayıldığın da ise (2.22) numaralı formüldeki gibi hesaplama yapılabilir (Nie ve ark., 1975; Lebe, 2006).

$$H(X) = -\sum_{i=1}^{I}\left(\frac{n_{i+}}{n}\right)\mathrm{In}\left(\frac{n_{i+}}{n}\right) \tag{2.18}$$

$$H(Y) = -\sum_{j=1}^{J}\left(\frac{n_{+j}}{n}\right)\mathrm{In}\left(\frac{n_{+j}}{n}\right) \tag{2.19}$$

$$H(XY) = -\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\frac{n_{ij}}{n}\right)\mathrm{In}\left(\frac{n_{ij}}{n}\right) \tag{2.20}$$

$$U_{(y/x)} = \frac{H(X)+H(Y)-H(XY)}{H(Y)} \tag{2.21}$$

$$U_{(x/y)} = \frac{H(X)+H(Y)-H(XY)}{H(X)} \tag{2.22}$$

Hangi değişkenin bağımlı değişken olduğunun önemli olmadığı durumlar için simetrik belirsizlik katsayısı kullanılır ve aşağıdaki gibi hesaplanabilir (Nehmzow, 2006; Anonim, 2020).

$$U_{(x,y)} = \frac{H(X)U_{X/Y}+H(Y)U_{Y/X}}{H(X)+H(Y)} = 2\left[\frac{H(X)+H(Y)-H(XY)}{H(X)+H(Y)}\right] \tag{2.23}$$

Şeklinde olup burada,

$U_{(y/x)}$: Y bağımlı, X bağımsız değişken olduğunda hesaplanan belirsizlik katsayısı,

$U_{(x/y)}$: X bağımlı, Y bağımsız değişken olduğunda hesaplanan belirsizlik katsayısı,
$U_{(x,y)}$: Simetrik belirsizlik katsayısıdır.

3.1.6 Cohen'in Kappa Katsayısı

Cohen'in Kappa katsayısı gerçek uyumluluk ölçütü olan ve tesadüfen beklenenin dışındaki, uyum olasılığının oranını gösteren katsayıdır. İsimsel ölçekte bir faktörün varlığını ya da yokluğunu belirlemek amacıyla uygulanarak uyumun değerlendirilmesine olanak sağlar (Sim ve Wright, 2005; Kılıç, 2009).

Cohen'in Kappa katsayısı -1 ve +1 arasında değerler alabilir. İki gözlemcinin arasında uyum olması durumunda +1 değerini alırken, -1 değerini aldığında uyumsuzdur. Cohen'in Kappa katsayısı 0 aldığında ise iki gözlemci arasındaki uyumun şansa bağlı olduğu bilinir (Dawson ve Trapp, 2004; Sim ve Wright, 2005; Kılıç, 2015).

Çizelge 3.4'de, d_{11}, değerlendiricilerin her ikisinin de "+" kategorisindeki sınıflandırma sayısını ve d_{00} ise değerlendiricilerin " − " kategorisindeki sınıflandırma sayısını gösterir. d_{01} ve d_{10}, her iki değerlendicinin de aynı fikirde olmadığı yani uyumsuz durumlardaki sınıflandırma sayısını gösterir. $d_{00}+d_{10}$ ve $d_{01}+d_{11}$, A değerlendiricisinin " − " ve "+" kategori içerisindeki; $d_{00}+d_{01}$ ve $d_{10}+d_{11}$, B değerlendiricisinin " − " ve "+" kategori içerisindeki sınıflandırma sayısını verir (Gwet, 2002; Temel ve Erdoğan, 2017).

Çizelge 3.4 İki Değerlendirici ve İki Kategoriye Ait Çapraz Tablo (Temel ve Erdoğan, 2017)

		Değerlendirici A		Toplam
		−	+	
Değerlendirici B	−	d_{00}	d_{01}	$d_{00}+ d_{01}$
	+	d_{10}	d_{11}	$d_{10}+ d_{11}$
Toplam		$d_{00}+d_{10}$	$d_{01}+d_{11}$	N

Cohen'in Kappa katsayısını bulmak için öncelikle toplam yani gözlenen uyum olasılığını ve Şansa bağlı uyum olasalığının bulunması gerekir. Aşağıdaki formüller yardımıyla hesaplanabilir (Machin ve ark., 2009; Temel ve Erdoğan, 2017).

$$P_a = \frac{d_{11} + d_{00}}{N} \tag{2.24}$$

$$P_e(\kappa) = \left[\left(\frac{(d_{00} + d_{10})}{N}\right)\left(\frac{(d_{00} + d_{01})}{N}\right)\right] + \left[\left(\frac{(d_{01} + d_{11})}{N}\right)\left(\frac{(d_{10} + d_{11})}{N}\right)\right] \tag{2.25}$$

Bulunan toplam ve şansa bağlı uyum olasılıkları kullanılarak Cohen'in Kappa katsayısı aşağıdaki gibi hesaplanır (Kundel ve Polansky, 2003; Gwet, 2008; Kanik ve ark., 2012).

$$\kappa = \frac{P_a - P_e(\kappa)}{1 - P_e(\kappa)} \tag{2.26}$$

Şeklinde olup burada,

P_a : Toplam uyum olasılığı,
$P_e(\kappa)$: Tesadüfe bağlı uyum olasılığı,
κ : Cohen'in Kappa katsayısıdır.

Cohen'in Kappa katsayısının yorumlanmaya uygun aralığı 0 ile +1 arasındadır. Güvenirlilik bakımından negatif (κ<0) değerlerin bir anlamı yoktur (Alpar, 2006; Gözükara Bağ ve ark., 2010).

Çizelge 3.5 Cohen'in Kappa Katsayısının Uyum Dereceleri (Landis ve Koch, 1977; Gözükara Bağ ve ark., 2010)

Dereceleri	Yorumları
≤0.20	Zayıf uyum
0.21 – 0.40	Ortanın altında uyum
0.41 – 0.60	Orta düzeyde uyum
0.61 – 0.80	İyi düzeyde uyum
0.81 – 1.00	Çok iyi düzeyde uyum

3.2 Sıralı (Ordinal) Değişkenler Arasındaki İlişkinin Belirlenmesinde Kullanılan İlişki Katsayıları

3.2.1 Goodman Kruskal'ın Gamma Katsayısı

Sıralama ölçekli verilerde kullanılabilen simetrik bir ilişki ölçüsü olan Gamma katsayısı, bütün veriler çapraz (kontenjans) tablodaki bir hücrede toplanmadığı sürece hesaplanabilir. Gamma katsayısı +1 ve -1 arasında değerler alır. Bu katsayı 0 olduğunda değişkenler arasında bağımsızlık vardır. Ayrıca 2×2'lik çapraz tabloları hariç bağımsızlık söz konusu olmadığında dahi Gamma katsayısı bazı zamanlarda 0 olabilir. Negatif ilişki -1 değerini aldığında ve +1'de ise pozitif ilişkiyi gösterir (Karagöz, 2010b; Dikmen, 2016).

Bir değişkendeki değişim pozitif ve diğer değişkendeki karşılık gelen değişim de pozitif olursa bu durum uyumluluk olarak adlandırılır. Bir değişkendeki değişim pozitif ve diğer değişkendeki karşılık gelen değişim negatif olduğunda ise bu uyumsuzluk olarak adlandırılır. Toplam uyumlu çift sayısı P ve toplam uyumsuz çift sayısı Q olmak üzere Gamma katsayısı aşağıdaki gibi bulunur (Öztuna ve ark., 2008).

$$G = \frac{P - Q}{P + Q} \tag{2.27}$$

Şeklinde olup burada,

G : Goodman Kruskal'ın Gamma katsayısı,
P : Toplam uyumlu çift sayısı,
Q : Toplam uyumsuz çift sayısıdır.

Gamma katsayı derecesinin istatistiksel olarak önemli olup olmadığının belirlenebilmesi için önem testi aşağıdaki gibi hesaplanabilir (İsrael, 2008).

$$Z = G\sqrt{\frac{P + Q}{N\left(1 - G^2\right)}} \tag{2.28}$$

Şeklinde olup burada,

N : Toplam gözlem sayısı,
Z : Z değeridir.

Goodman Kruskal'ın Gamma önemlilik testinde elde edilen Z değeri 1.96 (%5 anlamlılık düzeyi için) veya daha fazla ise, değişkenler arasında hiçbir ilişkinin bulunmadığını varsayan H_0 hipotezi reddedilir (İsrael, 2008).

3.2.2 Kendall Tau-b Katsayısı

Kendall tau-b ilişki katsayısı 2×2'lik çapraz tablolarda daha çok kullanılır. Bağımlı ve bağımsız değişkenlerdeki eşit sıra numaraları için düzeltme yapar. Karesel ya da karesel olmayan tablolarda istatistiksel bağımsızlık söz konusuysa

Kendall tau-b katsayısı 0 değerini alır. Yalnızca karesel tablolarda tüm değerler tek bir köşegende yerleşmişse +1 veya -1 değerini alma durumu olur (Öztuna ve ark., 2008).

Kendall tau-b katsayısı parametrik olmayan korelasyon katsayısıdır. Spearman rank katsayısına bir alternatiftir. İki sıralı değişken arasındaki ilişkilendirmenin bir ölçüsüdür ve bağlı sıraları dikkate alır. Çok sayıda bağlı sıraya sahip küçük veri setleri için kullanılabilir. Spearman rank katsayısı gibi Kendall tau-b katsayısı için de tüm veriler her değişken üzerinde sıralanır. Ancak, Pearson korelasyonu veya Spearman rank katsayılarından farklı bir şekilde hesaplanır. Kendall tau-b, ikinci değişken üzerindeki sıralama düzeninin birinci değişken üzerindeki sıralama düzeniyle ne kadar iyi eşleştiğini değerlendirir (Hinton ve ark., 2004).

Kendall tau-b olasılığı gösterir. İki değişkende aynı sırada yer alan (uyuşan) verilerin gözlenme olasılığı aynı sırada yer almayan (uyuşmayan) verilerin gözlenme olasılığı arasındaki farktır. Yani aynı sırada yer alan iki serideki veriler uyuştuğundan Kendall tau-b katsayısı +1 değerini alır. Farklı sırada yer alan serideki verilerde ise uyuşma olmadığından katsayının -1 değerini aldığı görülür (Terzi, 2019).

İki satır iki sütunlu çapraz tablolarından hesaplanan Kendall tau-b katsayısı, Pearson korelasyon katsayısına eşittir. Kendall tau-b katsayısı aşağıdaki gibi hesaplanır (Karagöz, 2010b; Dikmen, 2016).

$$T_b = \frac{2(P-Q)}{\sqrt{\left(N^2 - \sum_{i=1}^{I} n_{i+}^2\right)\left(N^2 - \sum_{j=1}^{J} n_{+j}^2\right)}} \tag{2.29}$$

Şeklinde olup burada,

T_b	:	Kendall tau-b katsayısı,
P	:	Uyumlu çiftlerin sayısı,
Q	:	Uyumsuz çiftlerin sayısı,
N	:	Toplam gözlem sayısı,
i	:	i. Satır,
j	:	j. Sütun,
n^2_{+j}		j. Sütun toplamının karesi,
n^2_{i+}		i. Satır toplamının karesidir.

Kendall tau-b katsayı derecesinin istatistiksel olarak önemli olup olmadığının belirlenebilmesi için önem testi aşağıdaki gibi hesaplanabilir (Üçkardeş, 2006).

$$Z = \frac{3T_b\sqrt{N(N-1)}}{\sqrt{2(2N+5)}} \tag{2.30}$$

Hesaplanan Z değeri 1.96 değerine eşit ya da bu değerden fazlaysa H_0 hipotezi reddedilerek değişkenler arasında istatistiksel olarak önemli bir ilişkinin bulunduğu sonucuna varılabilir.

3.2.3 Kendall Tau-c Katsayısı

Kendall tau-c testi, Kendall tau-b testinin uygulanamadığı durumlarda kullanılan bir testtir. Kendall tau-c testi, kare veya dikdörtgen şeklinde olan daha büyük tablolar için kullanılır. Kaynaklarda "Kendall Stuart'ın Tau c" ya da "Stuart'ın Tau c" adlarından biriyle de rastlanılabilir (Ergün, 1995; Üçkardeş, 2006).

Kendall tau-b yalnızca kare yani eşit sayıda satır ve sütun sayısı olan tablolar için +1 veya -1 değerlerine ulaşır. Kare olmayan tablolar için ise Kendall tau-c katsayısı kullanılması önerilir. Kendall tau-c büyük tablolar için ve özellikle kare olmayan beklenmedik tablolar için Kendall tau-b'nin bir modifikasyonudur (De Muth, 2006).

Kendall tau-c ilişki katsayısı -1 ve +1 arasında değişen değerler alır. Ayrıca bu aralık değerlerinde olabilmesi için özel bir şekilde düzenlenmiştir (Öztuna ve ark., 2008).

Kendall Tau değeri a, b ve c olarak üç şekilde hesaplanır. Farklı Tau değerlerinin hesaplanmasının sebebi pay kısmında aynı formül kullanılırken payda kısmında yer alan ifadenin farklı olmasıdır. Kendall tau-a, +1 ve -1 değerlerine veri setinde bağ olması sebebiyle ulaşamadığı görülür. Çünkü payda da yer alan değer pay kısmındaki değerden her zaman büyük olur. Bunun gibi dezavantajlarından dolayı Kendall tau-a yerine, Kendall tau-b ve Kendall tau-c, bağları dikkate aldığı için bu katsayılar kullanılır (Altaş ve ark., 2012).

Kendall tau-c katsayısı aşağıdaki gibi hesaplanır (Stuart, 1953).

$$T_c = \frac{2(P-Q)}{\frac{N^2(k-1)}{k}} = \frac{2k(P-Q)}{N^2(k-1)} \tag{2.31}$$

Şeklinde olup burada,

T_c : Kendall tau-c katsayısı,
N : Toplam gözlem sayısı,
$P-Q$: Uyumlu ve uyumsuz çiftlerin farkı,
k : Satır ya da sütun sayısından küçük olanı ifade eder.

Kendall tau-c katsayı derecesinin istatistiksel olarak önemli olup olmadığının belirlenebilmesi için aşağıdaki gibi H_0 ve H_1 hipotezleri kurularak Kendall tau-c önem testi aşağıdaki gibi hesaplanabilir (Terzi, 2019).

H_0: İlişki yoktur.

H_1: İlişki vardır.

$$Z = \frac{T_c}{\sqrt{\frac{2(2N+5)}{9N(N-1)}}} \sim Z_\alpha \tag{2.32}$$

Bulunan Z değeri, 1.96 değerinden büyük olduğunda H_0 hipotezi reddedilerek istatistiksel olarak değişkenler arasında önemli bir ilişkinin bulunduğuna ve Z değeri küçük ise H_0 hipotezi kabul edilerek istatistiksel olarak önemli bir ilişkinin olmadığına karar verilir.

3.2.4 Kendall Tau Katsayısı

Kendall Tau katsayısı, Spearman rank katsayısına alternatif olarak önerilir. Bu katsayı "Kendall sıralama katsayısı" olarak da isimlendirilir (Kendall, 1938).

Kendall Tau tahmin edicisi, uyumlu olan çiftlerin olasılığı ile uyumsuz olan çiftlerin olasılığı arasındaki fark bulunarak elde edilir (Diker, 2009; Kılıç Topal, 2015).

Değişkenler arasındaki uyum ilişkileri yorumuna dayanan Kendall tau basit yaygın olarak kullanılan bir ilişki katsayısıdır. İki çift gözlem (X_i, Y_i) ve (X_j, Y_j)'nin $X_i<X_j$ ve $Y_i<Y_j$ veya $X_i>X_j$ ve $Y_i>Y_j$ olması durumunda uyumlu olduğu söylenir. Eşdeğer olarak eğer çiftler $(X_i-X_j)(Y_i-Y_j)>0$ ise uyumludur. $(X_i-X_j)(Y_i-Y_j)<0$ olduğunda da çiftler uyumsuzdur (Yenigün, 2007).

Her ikisi de $X_i>X_j$ ve $Y_i<Y_j$ veya $X_i<X_j$ ve $Y_i>Y_j$ ise uyumsuz gözlemler çifti olduğu söylenebilir. Eğer $X_i=X_j$ ve/veya $Y_i=Y_j$, olduğunda ise gözlemler çifti ne uyumlu ne de uyumsuzdur (Puth ve ark., 2015).

. P, toplam uyumlu çiftlerin sayısı ile Q, toplam uyumsuz olan çiftlerin sayısı olmak üzere Kendall Tau katsayısı aşağıdaki gibi hesaplanabilir (Kılıç Topal, 2015).

$$r_T = \frac{P-Q}{\frac{n(n-1)}{2}} \tag{2.33}$$

Şeklinde olup burada,

r_T : Kendall Tau katsayısı,
P : Uyumlu çiftlerin sayısı,
Q : Uyumsuz çiftlerin sayısı,
n : Toplam gözlem değerleridir.

X_i ile X_j arasındaki fark Y_i ile Y_j arasındaki fark aynı doğrultuda değil ise çiftlerin uyumlu olmadığı sonucuna varılır. Kendall Tau katsayısı +1 ile -1 arasında değişen değerler alır. Bu katsayı, tüm çiftlerin uyumlu olması durumunda +1 ve uyumlu olmadığında ise -1'e eşit olur (Altaş ve ark., 2012).

3.2.5 Spearman Rank Katsayısı

Spearman rank katsayısı -1 ve +1 arasında değişen değerler alır. Bu katsayı -1 olduğunda negatif mükemmel bir doğrusal ilişkinin, +1 olduğunda ise bir değişkenin sıraları ile diğer değişkenin sıraları arasında pozitif mükemmel bir doğrusal ilişkinin varlığı söylenir. Spearman rank katsayı -1 ya da +1 olan bu iki değerden birine ne kadar fazla yakın olursa, ona göre o kadar güçlü bir doğrusal ilişkinin var olmasıyla bilinir. Katsayı değeri 0 olduğunda iki değişkenin sıraları arasında doğrusal bir ilişkinin olmadığı görülür (Akgül, 2003; Öztuna ve ark., 2008).

İki değişken arasındaki doğrusal ilişki ya da bir değişkenin iki ve daha fazlası ile olan ilişkisini; X ve Y değişkenlerinin gerçek gözlem değerleri değil bu değerlerin sıra numaraları göz önünde bulundurularak hesaplama yapılır. X değişkenindeki sıralamanın Y değişkenindeki sıralamaya nasıl bir uygunluk sağladığına bakılır. Bu katsayı parametrik olmayan bir korelasyon yöntemidir. Değişkenlerden biri sıralı ölçekle elde edilmiş ve en az biri normal dağılım göstermiyorsa bununla birlikte örneklem sayısı küçük ve iki değişken arasındaki ilişki de doğrusallıktan ayrılmış ise Spearman rank katsayısı kullanılabilir. Bu katsayı aşağıdaki gibi hesaplanır (Hamarat, 2017).

$$r_s = 1 - \frac{6\sum D_i^2}{n\left(n^2-1\right)} \tag{2.34}$$

şeklinde olup burada,

r_s : Spearman rank ilişki katsayısı,
n : Gözlem sayısı,
D_i : $X_i - Y_i$ olup, X_i ve Y_j çiftlerinin sıra numaraları arasındaki farktır.

Spearman rank katsayısının hipotez kontrolü, örnek hacmi 10'dan fazla ise Pearson korelasyon katsayısındaki gibi t değerini hesaplayarak yapılır. Genelde daha küçük örneklerle çalışıldığı için Spearman rank katsayısının önemlilik testinin yapılmasında örnek genişliği dikkate alınarak hazırlanmış kritik Spearman rank tablo değerinden yararlanılır. Örnek hacmi 10'dan fazla olması durumunda t değeri hasaplanır. Bu t değeri, (n-2) serbestlik dereceli t tablo değerine eşit veya büyük olduğu durumda sıfır hipotezi reddedilerek değişkenlerin arasında önemli bir bağlılığın ya da ilişkinin bulunduğu sonucuna varılır. İlişkinin var olması ile birlikte ilişkinin derecesi hesaplanabilir. t değeri, tablo değerinden küçük ise sıfır hipotezi kabul edilir. Yani önemli bir bağlılığın ya da ilişkinin olmadığını ifade eder. Bunun için hipotez kontrolü aşağıdaki formül yardımıyla hesaplanarak yapılır (Mendeş, 2012).

$$S_r = \sqrt{\frac{1-r_s^2}{n-2}} \tag{2.35}$$

$$t = \frac{r_s}{\sqrt{\frac{1-r_s^2}{n-2}}} \tag{2.36}$$

Şeklinde olup burada,

S_r : Standart hata,
t : t değeridir.

3.2.6 Somer'in d Katsayısı

Sıralı değişkenler arasındaki ilişkinin incelenmesi amacıyla hesaplanan diğer katsayılar bağımlı ve bağımsız değişken ayrımı yapmamasına rağmen Somer'in d katsayısı, değişkenlerden birisinin bağımlı değişken olarak belirlenebileceğini varsayar (Siegel ve Castellan, 1988; Öztuna ve ark., 2008). Somer'in d katsayısı sıfırdan farklı en az iki hücre frekansının satır ile sütunlarında var olması durumunda hesaplanır. Bu katsayı -1 ve +1 arasında değişen değerler alır. Bu katsayının 0 değerini alması iki değişkenin birbirinden bağımsız olduğunu gösterir. Bu durum 2×2 boyutunda olan tablolar için geçerli değildir (Goodman ve Kruskal, 1972).

Multinominal örnekleme modeline bağlı olarak, X bağımsız değişken ve Y bağımlı değişken olması durumunda $\hat{d}_{YX}$ katsayısının en yüksek olabilirlilik tahmincisi (2.37) numaralı formül yardımıyla hesaplanır. Benzer şekilde, Y bağımsız değişken ve X bağımlı değişken olması durumunda $\hat{d}_{XY}$ katsayısının en yüksek olabilirlilik tahmincisi (2.38) numaralı formül yardımıyla hesaplanır (Somers, 1962).

$$\hat{d}_{YX} = \frac{2(P-Q)}{N^2 - \sum_{i=1}^{I} n_{i+}^2} \tag{2.37}$$

$$\hat{d}_{XY} = \frac{2(P-Q)}{N^2 - \sum_{j=1}^{J} n_{+j}^2} \tag{2.38}$$

d katsayısının en yüksek olabilirlilik tahmincileri ise (2.39) numaralı formülü ile elde edilir (Somers, 1962).

$$d = \frac{4(P-Q)}{2N^2 - \sum_{i=1}^{I} n_{i+}^2 - \sum_{j=1}^{J} n_{+j}^2} \tag{2.39}$$

Şeklinde olup burada,

P : Uyumlu çiftlerin sayısı,
Q : Uyumsuz çiftlerin sayısı,
N : Toplam gözlem sayısı,
i : i. satır,
j : j. sütun,
n_{i+}^2 : i. satır toplamının karesi,
n_{+j}^2 : j. sütun toplamının karesidir.

Somer'in d katsayısının önem testi aşağıdaki gibi hesaplanabilir (Everitt, 2001; Üçkardeş, 2006).

$\hat{d}_{YX}$ için test istatistiği,

$$Z = \frac{\hat{d}_{YX}}{\sqrt{\frac{4\left(c^2-1\right)\left(r+1\right)}{9Nc^2\left(r-1\right)}}} \tag{2.40}$$

$\hat{d}_{XY}$ için test istatistiği,

$$Z = \frac{\hat{d}_{XY}}{\sqrt{\frac{4\left(r^2-1\right)\left(c+1\right)}{9Nr^2\left(c-1\right)}}} \tag{2.41}$$

Şeklinde olup burada,

c : Sütun sayısı,
r : Satır sayısı,
N : Toplam gözlem sayısı,
Z : Z değeridir.

Bulunan Z değeri, 1.96 değerinden büyük olduğunda H_0 hipotezi reddedilerek istatistiksel olarak değişkenler arasında önemli bir ilişkinin bulunduğuna ve Z değeri küçük ise H_0 hipotezi kabul edilerek istatistiksel olarak önemli bir ilişkinin olmadığına karar verilir.

3.3 İsimsel (Nominal) ve Kesikli/Sürekli Değişkenler Arasındaki İlişkinin Belirlenmesinde Kullanılan İlişki Katsayıları

3.3.1 Nokta Çift Serili Korelasyon Katsayısı

Nokta çift serili korelasyon katsayısı, -1 ile +1 arasında değişen değerler alır. Mükemmel negatif ilişkiyi -1 değeri, mükemmel bir pozitif ilişkiyi +1 değeri gösterir. Hiçbir ilişki olmaması durumunda katsayı 0 değerini alır. Tüm korelasyon katsayıları, nedensel bir ilişki ifade etmeyen bağımlılık ölçütleridir. Nokta çift serili korelasyon katsayısı, Pearson'un iki değişkenli korelasyon katsayısının hesaplanacağı şekilde hesaplanır. Burada iki değişkenin iki değişkenli değeri 0 ya da 1'dir. Bu nedenle binary (iki değerli) olarak da adlandırılır. Tüm korelasyon analizlerinde olduğu gibi nokta çift serili korelasyonda da iki değişken arasındaki ilişkinin veya birlikte oluşun gücünü ölçer. Korelasyon analizleri bu birliktelik gücünü tek bir değerde, korelasyon katsayısında ifade eder. Nokta çift serili korelasyon katsayısı, sürekli değişken (oran veya aralık verileri) ile iki değerli değişken arasındaki ilişki kuvvetinin korelasyon ölçüsüdür. İki değerli değişkenler, isimsel ölçeğin sadece iki değeri olan değişkenleridir (Anonim, 2019b).

Nokta çift serili korelasyon katsayısı aşağıdaki gibi hesaplanır (Büyüköztürk ve ark., 2000).

$$S_y = \sqrt{\frac{n\sum Y^2 - (\sum Y)^2}{n(n-1)}} \tag{2.42}$$

$$r_{pq} = \left(\frac{\bar{Y}_p - \bar{Y}_q}{S_y}\right)\sqrt{pq} \tag{2.43}$$

Şeklinde olup burada,

S_y : Sürekli değişkenin standart sapması,
r_{pq} : Nokta çift serili korelasyon katsayısı,
$\bar{Y}_p$, $\bar{Y}_q$: Süreksiz değişkendeki iki kategorinin sürekli değişkendeki ölçümleri ortalamaları,
p, q : Süreksiz değişkendeki iki kategorinin oranlarıdır.

Determinasyon katsayısı denilen belirlilik katsayısı ile bir bağımlı değişkendeki değişimin bir bağımsız değişkenle açıklanabilen değişkenlik oranıdır (Sen, 2012).

Nokta çift serili korelasyon katsayısının karesi alınarak etki eden faktörün ilişkili olduğu diğer faktörde yarattığı değişkenliğe belirlilik katsayısı denir. Determinasyon katsayısı olarak da adlandırdığımız bu katsayı aşağıdaki gibi hesaplanır (Şencan, 2005).

$$R^2 = (r_{pq})^2 \tag{2.44}$$

Şeklinde olup burada,

R^2 : Determinasyon katsayısı,
r_{pq} : Nokta çift serili korelasyon katsayısıdır.

Nokta çift serili korelasyon katsayısının önem kontrolü aşağıdaki gibi hesaplanır (Mendeş, 2012).

$$t = \frac{r_{pq}\sqrt{n-2}}{\sqrt{1-r_{pq}^{2}}} \qquad (2.45)$$

Bu şekilde hesaplanın t değeri (n-2) serbestlik dereceli t dağılımı gösterir. Hesaplanan t değeri, bulunan serbestlik dereceli t tablo değerinden küçük olduğunda H_0 hipotezi kabul edilerek önemli bir ilişkinin olmadığına, büyük olduğunda ise sıfır hipotezi reddedilerek önemli bir ilişkinin varlığına karar verilir.

3.3.2 Çift Serili Korelasyon Katsayısı

Çift serili korelasyon katsayısı, iki kategorili yapay süreksiz bir değişken ile en az aralık ölçeğinde ölçülen sürekli bir değişken arasındaki doğrusal bir ilişkiyi açıklamada kullanılır (Büyüköztürk ve ark., 2000).

Çift ve nokta serili korelasyon katsayıları sadece kavramsal olmayan bir farkla ayırt edilir, ancak istatistiksel hesaplamaları oldukça farklıdır. Bu korelasyon katsayıları, iki değişkenlerden biri iki kategorili olduğu durumda kullanılır. Çift ve nokta serili korelasyonların kullanımları arasındaki fark, ikili değişkenin kesikli veya sürekli olmasına bağlıdır. Nokta serili korelasyon katsayısı bir değişkenin ikili isimsel kesikli olduğu durumlarda kullanılırken, çift nokta serili korelasyon katsayısı bir değişkenin ikili isimsel sürekli olduğu zaman kullanılır. Çift nokta serili korelasyon katsayısı doğrudan SPSS'de hesaplanamaz. Önce nokta serili korelasyon katsayısını hesaplamanız ve daha sonra bu rakamı düzeltmek için bir denklem kullanmanız gerekir (Field, 2000).

SPSS ile nokta korelasyon katsayısı hesaplandıktan sonra gerekli olan düzeltme için kullanılacak denklem aşağıdaki gibidir (Büyüköztürk ve ark., 2000).

$$r_b = \frac{r_{pq}\sqrt{pq}}{h} \qquad (2.46)$$

Çift serili korelasyon katsayısı aşağıdaki gibi hesaplanır (Büyüköztürk ve ark., 2000).

$$r_b = \left(\frac{\bar{Y}_p - \bar{Y}_q}{S_y} \right) \frac{pq}{h} \tag{2.47}$$

Şeklinde olup burada,

r_b	: Çift serili korelasyon katsayısı,
r_{pq}	: Nokta çift serili korelasyon katsayısı,
S_y	: Sürekli değişkenin standart sapması,
$\bar{Y}_p, \bar{Y}_q$	: Yapay süreksiz hale getirilmiş değişkenlerdeki iki kategorinin sürekli değişkenlerdeki ölçümleri ortalamaları,
p, q	: Yapay süreksiz değişkendeki iki kategorinin oranları,
h	: Normal dağılım eğrisi altında kalan alanda p ve q’yu ayıran noktanın ordinat yüksekliğidir.

Çift serili korelasyon katsayısının önem testi aşağıdaki gibi hesaplanır (Büyüköztürk ve ark., 2000).

$$Z = \frac{hr_b}{\sqrt{\frac{pq}{n}}} \tag{2.48}$$

Z değeri, 1.96 değerinden büyük olduğunda H_0 hipotezi reddedilerek istatistiksel olarak değişkenler arasında önemli bir ilişkinin bulunduğuna ve Z değeri küçük ise H_0 hipotezi kabul edilerek istatistiksel olarak önemli bir ilişkinin bulunmadığına karar verilir.

3.4 İsimsel (Nominal) ve Aralık (İnterval) Değişkenler Arasındaki İlişkinin Belirlenmesinde Kullanılan İlişki Katsayıları

3.4.1 Eta Katsayısı

Değişkenler arasında doğrusal ilişkinin olmadığı durumda kullanılan katsayıdır. Yapılan korelasyon araştırmalarında hipotez yani öngörü, ilişkinin doğrusal olmasıdır. İki değişken olduğunda bazen ilişki doğrusal değildir. Bu duruma bağlı olarak Eta katsayısından yararlanmak uygundur (Borg, 2006; Shaldehi, 2013).

Araştırmacı Eta katsayısını, bir aralığa bağımlı değişken ve isimsel veya sıralı bağımsız bir değişkenle karşılaştığında dikkate almalıdır. Cramer’in V katsayısı gibi sadece 0 ve +1 arasında değişebilen değerler alır. Phi ve Cramer V'nin etkinleştirilmesini sağlayarak çapraz tablolor yöntemi ile hesaplanabilir. Eta katsayısı ilişkinin gücünü belirlemeyi sağlayan yararlı bir test olarak kabul edilir (Bryman ve Cramer, 2005).

Uygulamada, Eta katsayısı hesaplanırken oransal ölçekli değişken, bağımlı değişken olarak ve isimsel ölçekli değişkenin ise bağımsız değişken olarak alınması gerekir. Bunun önemli bir sonucu analize, diğer bağımsız değişkenler dahil edildiğinde Eta katsayısının karşılaştırılabileceği çok değişkenli bir katsayı yoktur. O zaman grup seçimi ile bağımsız değişkenler arasındaki iki değişkenli korelasyonu Eta katsayısı için oransal ölçek olarak incelenmesi yararlıdır (Knutsen, 2017).

Çizelge 3.6 Karşılaştırılacak Grup Sayısının n Tekerrürlü Denendiği Bir Denemeden Elde Edilen Verilerin Özetlenmesi (Kesici ve Kocabaş, 2007)

	1	**2**	**3…**	**k**	
	x_{11}	x_{21}	x_{31}	x_{k1}	
	x_{12}	x_{22}	x_{32}	x_{k2}	
	x_{13}	x_{23}	x_{33}	x_{k3}	
	…	…	…	…	
	x_{1n}	x_{2n}	x_{3n}	x_{kn}	
Toplam	T_1	T_2	T_3	T_k	T
Ortalama	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_k$	$\bar{\bar{x}}$

Gruplar arası kareler toplamı (2.49) numaralı ve genel kareler toplamı (2.50) numaralı formüller ile aşağıdaki gibi hesaplanabilir (Kesici ve Kocabaş, 2007).

$$\mathrm{GAKT} = \sum_{i=1}^{k} n_i (\bar{x}_i - \bar{\bar{x}})^2 \tag{2.49}$$

$$\mathrm{GKT} = \sum_{i=1}^{k} \sum_{j=1}^{n} (x_{ij} - \bar{\bar{x}})^2 \tag{2.50}$$

Şeklinde olup burada,

n : Örneklem genişliği,
k : Karşılaştırılacak grup sayısı,
n_i : i. gruptaki gözlem sayısı,
x_{ij} : i. muamele grubundaki j. tekerrüre ait gözlem değeri,

$\bar{x}_i$: i. grubun ortalaması,
$\bar{\bar{x}}$: Ortalamaların ortalamasıdır.

Ayrıca Eta (η_{YX}) katsayısı; Y değişkeni bağımlı değişken olarak varsayıldığında gruplar arası kareler toplamı (2.51) numaralı ve genel kareler toplamı (2.52) numaralı formüller ile hesaplanır. X değişkeni bağımlı değişken olarak varsayıldığında Eta (η_{XY}) katsayısı gruplar arası kareler toplam (2.54) numaralı ve genel kareler toplamı (2.55) numaralı aşağıdaki formüller yardımıyla hesaplanabilir (Mueller ve ark., 1970; Lebe, 2006).

$$GAKT_{YX} = \sum_{i=1}^{I} \frac{\left[\sum_{j=1}^{J}(j)(n_j)\right]^2}{n_{i+}} - \frac{\left[\sum_{i=1}^{I}\sum_{j=1}^{J}(j)(n_j)\right]^2}{n} \tag{2.51}$$

$$GKT_{YX} = \sum_{j=1}^{J}(j^2)(n_{+j}) - \frac{\left[\sum_{i=1}^{I}\sum_{j=1}^{J}(j)(n_j)\right]^2}{n} \tag{2.52}$$

$$\eta_{YX} = \sqrt{\frac{GAKT_{YX}}{GKT_{YX}}} \tag{2.53}$$

$$GAKT_{XY} = \sum_{j=1}^{J} \frac{\left[\sum_{i=1}^{I}(i)(n_i)\right]^2}{n_{j+}} - \frac{\left[\sum_{j=1}^{J}\sum_{i=1}^{I}(i)(n_i)\right]^2}{n} \tag{2.54}$$

$$GKT_{XY} = \sum_{i=1}^{I}(i^2)(n_{i+}) - \frac{\left[\sum_{j=1}^{J}\sum_{i=1}^{I}(i)(n_i)\right]^2}{n} \tag{2.55}$$

$$\eta_{XY} = \sqrt{\frac{GAKT_{XY}}{GKT_{XY}}} \tag{2.56}$$

Şeklinde olup burada,

n : Örneklem genişliği,
i : i. satır,
j : j. sütun,
GAKT : Gruplar arası kareler toplamı,
GKT : Genel kareler toplamı,
η : Eta katsayısıdır.

Eta katsayısının anlamlılığını test etmek için F testinden yararlanılır. İstatistiksel olarak önemli olup olmadığının belirlenebilmesi için önem testi aşağıdaki gibi hesaplanabilir (Bruning ve Kintz, 1993).

$$GIKT = GKT - GAKT \tag{2.57}$$

$$GASD = k - 1 \tag{2.58}$$

$$GISD = N - k \tag{2.59}$$

$$GAKO = \frac{GAKT}{GASD} \tag{2.60}$$

$$GIKO = \frac{GIKT}{GISD} \tag{2.61}$$

$$F = \frac{GAKO}{GIKO} \tag{2.62}$$

Şeklinde olup burada,

GIKT : Gruplar içi kareler toplamı,
GASD : Gruplar arası serbestlik derecesi,
GISD : Gruplar içi serbestlik derecesi,
GAKO : Gruplar arası kareler ortalaması,
GIKO : Gruplar içi kareler ortalaması,
N : Örnek genişliği,
k : Karşılaştırılacak grup sayısı,
F : F değeridir.

Hesaplanan F değeri, gruplar içi serbetlik derecesi ve gruplar arası serbestlik derecesine karşılık gelen F tablo değerine eşit ya da F tablo değerinden büyük ise H_0 hipotezi reddedilir (Mendeş, 2012).

4. MATERYAL ve YÖNTEM

4.1 Materyal

Çalışmanın materyalini MSDEV ve IMSL Library desteği ile Fortran programlama dilinde yazılmış simülasyon programı ile iki değişkenli normal dağılımdan çekilen tesadüf sayıları oluşturmaktadır. Aralarında sırasıyla 0.5 ve 0.9 korelasyon bulunan iki tesadüf değişkeninin örneklem genişlikleri sırasıyla 30, 50, 100, 150 ve 200 olarak belirlenmiştir.

4.2 Yöntem

4.2.1 Simülasyon Çalışması

Bu tez çalışmasında bir simülasyon programı hazırlanmış ve bulgular 50.000 simülasyon denemesi sonucunda elde edilmiştir. İlk olarak aralarında sırasıyla 0.5 ve 0.9 korelasyon bulunan ve örneklem genişlikleri sırasıyla 30, 50, 100, 150 ve 200 olan iki değişkenli normal dağılımdan iki tesadüf değişkeni üretilmiştir. Bu değişkenlere ait tesadüf sayıları eşit aralıklı olarak bölünmüş ve sırasıyla kontenjans tablo boyutu 3×3, 4×4 ve 5×5 ölçekli olacak şekilde kodlanmıştır. Belirlenen populasyon ilişki düzeyleri, örneklem genişlikleri ve tablo boyutları için Pearson, Spearman rank, Kendall tau-b, Kendall tau-c, Goodman Kruskal'ın Gamma ve Somer'in d katsayıları hesaplanmıştır. 50.000 simülasyon denemesi sonucunda hesaplanan bu katsayıların ortalamaları alınmış ve performansları birbirleri ile karşılaştırılmıştır. Karşılaştırma kriteri olarak populasyondaki gerçek ilişki düzeyine yakınlık düzeyi kullanılmıştır.

Çalışmada dikkate alınan deneme koşulları Çizelge 4.1'de verilmiştir.

Çizelge 4.1 Dikkate Alınan Deneme Koşulları

İlişki katsayısı	Pearson, Spearman rank, Kendall tau-b, Kendall tau-c, Gamma, Somer'in d
Tablo boyutu	3×3, 4×4, 5×5
Örneklem genişliği (n)	30, 50, 100, 150, 200
Populasyondaki Korelasyon Katsayısı (ρ)	0.50, 0.90
Deneme sayısı	50.000 simülasyon denemesi

4.2.2 İncelenen İlişki Katsayıları

4.2.2.1 Pearson Korelasyon Katsayısı

Aralık ölçeklerle elde edilmiş, x ve y değişkenlerin ilişkisinin gücünü ve yönünü araştırmak amacıyla kullanılan test Pearson Korelasyon katsayısıdır. Bu katsayı, x ve y değişkeni arasındaki verilerin aralık ölçekle alındığında, değişkenlerin ikisi de normal dağılımdan gelerek popülasyondan tesadüfi şekilde seçilmesiyle kullanılır (Üçkardeş, 2006).

n gözlem bulunan iki değişkenli bir veri seti için, $(x_1, y_1), (x_2, y_2), \ ...(x_n, y_n)$ Pearson Korelasyon katsayısı aşağıdaki gibi hesaplanır (Tuğran, 2015).

$$r_{xy} = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}} = \frac{\sum d_x d_y}{\sqrt{(\sum d_x^2)\left(\sum d_y^2\right)}} \tag{2.63}$$

Şeklinde olup burada,

r_{xy}	:	Pearson Korelasyon katsayısı,
$\sum d_x d_y$	:	Çarpımlar toplamı,
$\sum d_x^2$	:	X özelliklerine ilişkin kareler toplamı,
$\sum d_y^2$	:	Y özelliklerine ilişkin kareler toplamıdır.

Bu şekilde hesaplanan korelasyon katsayısı ya da doğrusal ilişkinin derecesinin istatistiksel olarak önemli olup olmadığının belirlenmesi amacıyla gerekli olan hipotez takımı:

H_0: $\rho_{xy} = 0$

H_1: $\rho_{xy} \neq 0$ şeklinde kurulur. H_0 hipotezinin test edilmesi için gerekli olan test istatistiği ise;

$$t = \frac{r_{xy}}{\sqrt{\frac{1 - r_{xy}^2}{n - 2}}} \tag{2.64}$$

Pearson korelasyon katsayısının yine bu katsayıya ilişkin standart hataya bölünmesiyle t testi bulunur. Bu şekilde hesaplanan t test istatistiği (n-2) serbestlik dereceli t dağılımı gösterir. Eğer bu şekilde hesaplanan t değeri, (n-2) serbestlik dereceli t tablo değerine eşit ya da büyük ise H_0 hipotezi reddedilir. Aksi durumda ise kabul edilir (Tuğran, 2015).

Çizelge 4.2 Pearson Korelasyon Katsayısının İlişki Kuvvet Dereceleri (Aslan ve ark., 2019)

Derece	**Yorum**
0.00 – 0.19	İlişki yok ya da önemsenmeyecek düzeyde ilişki
0.20 – 0.39	Zayıf ilişki
0.40 – 0.69	Orta düzeyde ilişki
0.70 – 0.89	Yüksek ilişki
0.90 – 1.00	Çok yüksek ilişki

4.2.2.2 Spearman Rank Katsayısı

Spearman rank katsayısı aşağıdaki gibi hesaplanır (Hamarat, 2017).

$$r_s = 1 - \frac{6\sum D_i^2}{n\left(n^2 - 1\right)} \tag{2.65}$$

şeklinde olup burada,

r_s : Spearman rank ilişki katsayısı,
n : Gözlem sayısı,
D_i : $X_i - Y_i$ olup, X_i ve Y_j çiftlerinin sıra numaraları arasındaki farktır.

Spearman rank katsayısının hipotez kontrolü aşağıdaki formül yardımıyla hesaplanarak yapılır (Mendeş, 2012).

$$S_r = \sqrt{\frac{1 - r_s^2}{n - 2}} \tag{2.66}$$

$$t = \frac{r_s}{\sqrt{\frac{1 - r_s^2}{n - 2}}} \tag{2.67}$$

Şeklinde olup burada,

S_r : Standart hata,
t : t değeridir.

4.2.2.3 Kendall Tau-b Katsayısı

Kendall tau-b katsayısı aşağıdaki gibi hesaplanır (Karagöz, 2010b; Dikmen, 2016).

$$T_b = \frac{2(P-Q)}{\sqrt{\left(N^2 - \sum_{i=1}^{I} n_{i+}^2\right)\left(N^2 - \sum_{j=1}^{J} n_{+j}^2\right)}} \tag{2.68}$$

Şeklinde olup burada,

T_b : Kendall tau-b katsayısı,
P : Uyumlu çiftlerin sayısı,
Q : Uyumsuz çiftlerin sayısı,
N : Toplam gözlem sayısı,
i : i. Satır,
j : j. Sütun,
n^2_{+j} j. Sütun toplamının karesi,
n^2_{i+} i. Satır toplamının karesidir.

Kendall tau-b katsayı derecesinin istatistiksel olarak önemli olup olmadığının belirlenebilmesi için önem testi aşağıdaki gibi hesaplanabilir (Üçkardeş, 2006).

$$Z = \frac{3T_b\sqrt{N(N-1)}}{\sqrt{2(2N+5)}} \tag{2.69}$$

Hesaplanan Z değeri 1.96 değerine eşit ya da bu değerden fazlaysa H_0 hipotezi reddedilerek değişkenler arasında istatistiksel olarak önemli bir ilişkinin bulunduğu sonucuna varılabilir.

4.2.2.4 Kendall Tau-c Katsayısı

Kendall tau-c katsayısı aşağıdaki gibi hesaplanır (Stuart, 1953).

$$T_c = \frac{2(P-Q)}{\frac{N^2(k-1)}{k}} = \frac{2k(P-Q)}{N^2(k-1)} \tag{2.70}$$

Şeklinde olup burada,

T_c : Kendall tau-c katsayısı,
N : Toplam gözlem sayısı,
$P-Q$: Uyumlu ve uyumsuz çiftlerin farkı,
k : Satır ya da sütun sayısından küçük olanı ifade eder.

Kendall tau-c katsayı derecesinin istatistiksel olarak önemli olup olmadığının belirlenebilmesi için aşağıdaki gibi H_0 ve H_1 hipotezleri kurularak Kendall tau-c önem testi aşağıdaki gibi hesaplanabilir (Terzi, 2019).

H_0: İlişki yoktur.

H_1: İlişki vardır.

$$Z = \frac{T_c}{\sqrt{\frac{2(2N+5)}{9N(N-1)}}} \sim Z_\alpha \tag{2.71}$$

Bulunan Z değeri, 1.96 değerinden büyük olduğunda H_0 hipotezi reddedilerek istatistiksel olarak değişkenler arasında önemli bir ilişkinin bulunduğuna ve Z değeri küçük ise H_0 hipotezi kabul edilerek istatistiksel olarak önemli bir ilişkinin olmadığına karar verilir.

4.2.2.5 Goodman Kruskal'ın Gamma Katsayısı

Toplam uyumlu çift sayısı P ve toplam uyumsuz çift sayısı Q olmak üzere Gamma katsayısı aşağıdaki gibi bulunur (Öztuna ve ark., 2008).

$$G = \frac{P-Q}{P+Q} \tag{2.72}$$

Şeklinde olup burada,

G : Goodman Kruskal'ın Gamma katsayısı,
P : Toplam uyumlu çift sayısı,
Q : Toplam uyumsuz çift sayısıdır.

Gamma katsayı derecesinin istatistiksel olarak önemli olup olmadığının belirlenebilmesi için önem testi aşağıdaki gibi hesaplanabilir (İsrael, 2008).

$$Z = G\sqrt{\frac{P+Q}{N(1-G^2)}} \tag{2.73}$$

Şeklinde olup burada,

N : Toplam gözlem sayısı,
Z : Z değeridir.

4.2.2.6 Somer'in d Katsayısı

Multinominal örnekleme modeline bağlı olarak, X bağımsız değişken ve Y bağımlı değişken olması durumunda $\hat{d}_{YX}$ katsayısının en yüksek olabilirlilik tahmincisi (2.74) numaralı formül yardımıyla hesaplanır. Benzer şekilde, Y bağımsız değişken ve X bağımlı değişken olması durumunda $\hat{d}_{XY}$ katsayısının en yüksek olabilirlilik tahmincisi (2.75) numaralı formül yardımıyla hesaplanır (Somers, 1962).

$$\hat{d}_{YX} = \frac{2(P-Q)}{N^2 - \sum_{i=1}^{I} n_{i+}^2} \tag{2.74}$$

$$\hat{d}_{XY} = \frac{2(P-Q)}{N^2 - \sum_{j=1}^{J} n_{+j}^2} \tag{2.75}$$

d katsayısının en yüksek olabilirlilik tahmincileri ise (2.76) numaralı formülü ile elde edilir (Somers, 1962).

$$d = \frac{4(P-Q)}{2N^2 - \sum_{i=1}^{I} n_{i+}^2 - \sum_{j=1}^{J} n_{+j}^2} \tag{2.76}$$

Şeklinde olup burada,

P : Uyumlu çiftlerin sayısı,
Q : Uyumsuz çiftlerin sayısı,
N : Toplam gözlem sayısı,
n_{i+}^2 : i. satır toplamının karesi,
n_{+j}^2 : j. sütun toplamının karesi,
i : i. satır,
j : j. sütundur.

Somer'in d katsayısının önem testi aşağıdaki gibi hesaplanabilir (Everitt, 2001; Üçkardeş, 2006).

$\hat{d}_{YX}$ için test istatistiği,

$$Z = \frac{\hat{d}_{YX}}{\sqrt{\frac{4\left(c^2-1\right)\left(r+1\right)}{9Nc^2\left(r-1\right)}}} \tag{2.77}$$

$\hat{d}_{XY}$ için test istatistiği,

$$Z = \frac{\hat{d}_{XY}}{\sqrt{\frac{4\left(r^2-1\right)\left(c+1\right)}{9Nr^2\left(c-1\right)}}} \tag{2.78}$$

Şeklinde olup burada,

c : Sütun sayısı,
r : Satır sayısı,
N : Toplam gözlem sayısı,
Z : Z değeridir.

Bulunan Z değeri, 1.96 değerinden büyük olduğunda H_0 hipotezi reddedilerek istatistiksel olarak değişkenler arasında önemli bir ilişkinin bulunduğuna ve Z değeri küçük ise H_0 hipotezi kabul edilerek istatistiksel olarak önemli bir ilişkinin olmadığına karar verilir.

5. BULGULAR ve TARTIŞMA

3×3 çapraz tablo düzeninde hesaplanan ilişki katsayıları Çizelge 5.1'de verilmiştir. Görüldüğü üzere tablo boyutu 3×3 için Pearson, Spearman rank, Kendall tau-b, Kendall tau-c, Somer'in d katsayılarına kıyasla Gamma ilişki katsayısının; popülasyonda ρ=0.5 olarak belirlenen gerçek ilişki derecesine daha fazla yaklaştığı görülmektedir. Gamma katsayısının örneklem genişliği arttıkça populasyonda belirlenen gerçek ilişki derecesinin bir miktar üzerinde sonuç verdiği görülmektedir. Kendall tau-b, Kendall tau-c, Somer'in d örneklem genişliği artsa dahi belirlenen gerçek ilişki derecesinden bağımsız olarak daha düşük sonuçlar verdiği görülmektedir.

Çizelge 5.1 3×3 Çapraz Tablo için Hesaplanan İlişki Katsayıları

ρ	İlişki Katsayısı	n				
		30	50	100	150	200
ρ=0.5	Pearson	0.406	0.400	0.398	0.396	0.396
	Spearman rank	0.405	0.400	0.398	0.396	0.395
	Kendall tau-b	0.364	0.367	0.370	0.371	0.369
	Kendall tau-c	0.365	0.367	0.370	0.371	0.369
	Gamma	0.481	0.511	0.550	0.569	0.580
	Somer'in d	0.364	0.367	0.370	0.371	0.369
ρ=0.9	Pearson	0.736	0.735	0.732	0.731	0.729
	Spearman rank	0.736	0.735	0.732	0.730	0.728
	Kendall tau-b	0.715	0.714	0.712	0.709	0.710
	Kendall tau-c	0.621	0.619	0.589	0.551	0.545
	Gamma	0.915	0.920	0.929	0.932	0.933
	Somer'in d	0.715	0.714	0.712	0.709	0.710

Çizelge 5.1'de, Gamma katsayısından sonra Pearson, Spearman rank ilişki katsayılarının da populasyondan beklenen gerçek ilişki derecesine biraz yaklaşsa da altında kaldığı ve Kendall tau-b, Kendall tau-c, Somer'in d katsayıları ise beklenen gerçek ilişki derecesinden daha düşük ve benzer sonuçlar verdiği görülmektedir. Fakat belirlenen ρ=0.9 ilişki düzeyinde ise Kendall tau-b, Somer'in d katsayılarının benzer sonuçlar bulduğu görülürken, Kendall tau-c bu iki katsayıdan farklı ve bu katsayıların da altında olup, belirlenen gerçek ilişki düzeyinden uzaklaşmaktadır. Populasyonda belirlenen gerçek ilişki derecesinin ρ=0.5'den ρ=0.9'a yükseltilmesi ile elde edilen sonuçlarda bunun dışında herhangi bir fark yaratmayacağı anlaşıldığı gibi Gamma katsayısının en iyi belirlenen gerçek ilişki derecesine yaklaştığı ve Kendall tau-b, Kendall tau-c, Somer'in d'nin uzaklaştığı yani belirlenen ilişki düzeyinin altında sonuç

verdiği görülmektedir. Örneklem genişliklerinin artmasıyla; Pearson korelasyon ve Spearman rank katsayılarının populasyonda belirlenen hem ρ=0.5 hem de ρ=0.9 için gerçek ilişki derecelerinden uzaklaştığı görülmektedir.

Çizelge 5.2'de yer alan 4×4 tablo boyutu için elde edilen bulgular incelendiğinde, tüm katsayıların 3×3 tablo boyutu için elde edilen sonuçlara oranla daha fazla populasyonun gerçek ilişki derecesine yaklaştığını göstermektedir. Gamma katsayısının, belirlenen ilişki derecesinin ρ=0.5 ve örneklem genişliğinin 100; ρ=0.9 için de örneklem genişliğinin 50 olduğu durumda populasyonda belirlenen gerçek ilişki derecesine en yakın sonuçları verdiğini göstermektedir. Örneklem genişliğinde azalış yada artış bir miktar populasyondaki belirlenen gerçek ilişki değerinden uzaklaştırsa da diğer katsayılara göre daha yakın sonuçlar verdiği görülmektedir. Farklı örneklem genişliklerinde Kendall tau-b ve Somer'in d, hem ρ=0.5 hem de ρ=0.9 olarak belirlenen gerçek ilişki derecelerine örneklem genişlikleri artsa bile tam olarak ulaşamamaktadır. Belirlenen gerçek ilişki derecelerine göre bu iki katsayının yaklaşık olarak aynı sonuçları verdiği görülmektedir.

Çizelge 5.2 4×4 Çapraz Tablo için Hesaplanan İlişki Katsayıları

ρ	**İlişki Katsayısı**	**n**				
		30	**50**	**100**	**150**	**200**
ρ=0.5	Pearson	0.431	0.432	0.434	0.434	0.431
	Spearman rank	0.430	0.431	0.432	0.433	0.430
	Kendall tau-b	0.361	0.378	0.380	0.382	0.386
	Kendall tau-c	0.342	0.340	0.339	0.330	0.327
	Gamma	0.464	0.470	0.500	0.519	0.520
	Somer'in d	0.361	0.378	0.380	0.382	0.386
ρ=0.9	Pearson	0.802	0.800	0.798	0.793	0.797
	Spearman rank	0.802	0.800	0.799	0.793	0.797
	Kendall tau-b	0.730	0.732	0.735	0.735	0.736
	Kendall tau-c	0.661	0.655	0.635	0.624	0.620
	Gamma	0.899	0.900	0.910	0.915	0.920
	Somer'in d	0.730	0.732	0.735	0.735	0.736

Tablo boyutu 4x4 olduğunda, belirlenen ρ=0.5 gerçek ilişki derecesi için Pearson ve Spearman rank katsayıları 30, 50, 100, 150 örneklem genişliklerinde artış göstermesine rağmen n=200 için bir miktar düştüğü; aynı şekilde ρ=0.9 için örneklem genişliğinin 30 ve 50 olduğu durum değerlendirildiğinde bu katsayıların derecesinde artış varken 100 ve 150 olan örneklem genişliklerinde bir miktar düştüğü görülmektedir. Pearson ve Spearman ilişki katsayıları örneklem genişliğinin artmasından olumsuz etkilenip populasyonda belirlenen gerçek ilişki derecesinden uzaklaşmasına rağmen Gamma katsayısından sonra gerçek ilişki derecesine yakın sonuç veren en iyi ilişki katsayıları olduğu görülmektedir. Kendall tau-c, örneklem genişliği artsa dahi populasyonda belirlenen ilişki derecesine en uzak olan birinci ilişki katsayısı olup ikincisi Kendall tau-b olduğu görülmektedir. Somer'in d ve Kendall tau b ilişki katsayılarının belirlenen ρ=0.5 ve ρ=0.9 gerçek ilişki dereceleri için örneklem genişliğinin 30, 50, 100, 150, 200 olduğu 3×3, 4×4 tablo boyutlarında benzer sonuçlar verdiği gözlemlenmektedir.

Çizelge 5.3'de tablo boyutu 5×5 için hesaplanan ilişki katsayıları incelendiğinde tablo boyutundaki artışın populasyondaki gerçek ilişki derecesine yaklaşmak için olumlu bir etkiye sahip olduğu görülmektedir. Belirlenen ilişki derecesinin ρ=0.5 ve örneklem genişliğinin 50 olduğu durumda populasyonda belirlenen gerçek ilişki derecesine Gamma katsayının tam olarak ulaştığı görülmektedir. Gamma katsayısı ρ=0.9 için tam populasyon değerinde olmasa bile yine en yakın sonucu vermektedir. Gamma katsayısı hem ρ=0.5 için hem de ρ=0.9 için örneklem genişliğinde artışa bağlı olarak gerçek ilişki derecesinin üzerinde sonuçlar verdiği görülmektedir. Pearson, Spearman rank ve Kendall tau-c katsayılarında da örneklem genişliğinin artması hesaplanan katsayıların populasyondaki gerçek değerden bir miktar uzaklaşmasına ve daha düşük değerler elde edilmesine sebep olmaktadır.

Çizelge 5.3 5×5 Çapraz Tablo için Hesaplanan İlişki Katsayıları

ρ	**İlişki Katsayısı**	**n**				
		30	**50**	**100**	**150**	**200**
ρ=0.5	Pearson	0.450	0.452	0.451	0.440	0.438
	Spearman rank	0.451	0.452	0.450	0.440	0.438
	Kendall tau-b	0.480	0.487	0.490	0.485	0.484
	Kendall tau-c	0.350	0.349	0.348	0.340	0.331
	Gamma	0.485	0.500	0.519	0.521	0.530
	Somer'in d	0.480	0.487	0.490	0.485	0.484
ρ=0.9	Pearson	0.835	0.830	0.820	0.805	0.800
	Spearman rank	0.834	0.830	0.820	0.805	0.801
	Kendall tau-b	0.770	0.765	0.759	0.749	0.748
	Kendall tau-c	0.710	0.700	0.689	0.670	0.651
	Gamma	0.920	0.923	0.930	0.940	0.940
	Somer'in d	0.770	0.765	0.760	0.749	0.748

Toblo büyüklüğü 5×5 olan, farklı örneklem genişliklerinde Kendall tau-b ve Somer'in d, hem ρ=0.5 hem de ρ=0.9 olarak belirlenen gerçek ilişki derecelerine, örneklem genişliklerinin artırılmasına rağmen ulaşamamakla birlikte bu iki katsayının benzer olarak aynı sonuçları elde ettiği görülmektedir. Gamma katsayısından sonra ρ=0.5 için gerçek ilişki derecesine en yakın sonuç veren Kendall tau-b ve Somer'in d iken ρ=0.9 için ise Pearson ve Spearman rank ilişki katsayıları olduğu anlaşılmaktadır. Belirlenen gerçek ilişki derecelerine göre Kendall tau-b ve Somer'in d katsayıları ρ=0.5 için örneklem genişliği 100 olduğunda gerçek ilişki seviyesine en yakın sonucu elde ederken ρ=0.9 için örneklem genişliği 30 olduğunda diğer örneklem genişliklerine göre daha iyi gerçek ilişki seviyesine yaklaştığı gözlemlenmektedir. Yani aynı tablo büyüklüğünde populasyonda belirlenen ilişki derecesinin yükseltilmesiyle daha az örneklem genişliğinde bu gerçek ilişki düzeyine yaklaşma eğilimi gösterdiği anlaşılmaktadır. Pearson ve Spearman rank ilişki katsayıları, örneklem genişliğinin artmasıyla orantılı olarak populasyonda belirlenen gerçek ilişki derecesinden uzaklaşma eğilimi göstermektedir. Kendall tau-c, örneklem genişliği artsa dahi populasyonda belirlenen ρ=0.5 ve ρ=0.9 için tüm örneklem genişliklerinde, gerçek ilişki derecesinden en çok uzaklaşan ilişki katsayısı olmaktadır.

Göktaş ve İşçi (2011), yaptıkları çalışmada Kendall tau-b, Kendall tau-c ve Somer'in d katsayıları; tablo boyutu arttıkça artmasına rağmen bu ilişki katsayıları her zaman belirlenen ilişki düzeylerinin altında sonuç verdiklerini, Gamma katsayısı tablo boyutu küçük ise, küçük örnek genişliği için iyi olduğunu ayrıca herhangi bir tablo boyutu için örnek büyüklüğü arttıkça Gamma katsayısının da arttığını ve belirlenen ilişki düzeylerinin üzerinde sonuç verdiğini bildirmişlerdir. Gamma katsayısı kare tablolar için gerçek ilişki derecesine yaklaştığını ve tablo boyutu yeterince büyük olduğunda, Pearson korelasyonu ve Spearman rank katsayısı, örneklem genişliğinden bağımsız olarak gerçek ilişki derecesine daha yakın, ancak Gamma katsayısı dışındaki sıralı katsayıların her zaman gerçek ilişki düzeyinin altında değer aldığı sonucuna varmışlardır.

Altaş ve ark., (2012) yaptıkları çalışmada tablo boyutu arttıkça Spearman rank katsayısı gerçek ilişki düzeyine yakın sonuç verdiğini ve değişkenler arasında ilişkinin ölçülmesinde daha tutarlı sonuç vermesinden dolayı bu katsayının kullanılmasının uygun olacağını belirtmişlerdir.

6. SONUÇ ve ÖNERİLER

Bu tez çalışmasında kategorik değişkenler arasındaki ilişkilerin belirlenmesi amacıyla kullanılan ilişki katsayıları incelenmiştir. Araştırıcıların hangi durumda hangi ilişki katsayısını hesaplamaları gerektiği hususu açıklanmaya çalışılırken ayrıca bazı korelasyon katsayılarının belirlenen deneme koşullarındaki performansları simülasyon yaklaşımı ile karşılaştırılmıştır.

50.000 simülasyon denemesi sonucunda; küçük boyutlu tablolarda örneklem genişliği arttıkça, Gamma katsayısı tüm tablo boyutlarında diğer tüm ilişki katsayılarından populasyondaki gerçek ilişki düzeyine daha yakın sonuçlar vermiştir. Bu durum örneklem genişliği arttıkça daha belirgin hale gelmiştir. Tablo boyutu arttıkça Gamma gerçek ilişki düzeyine yaklaşmıştır.

Kendall tau-b ve Kendall tau-c, tüm deneme koşullarında gerçek ilişki düzeyinin altında sonuçlar vermiştir. Belirlenen tüm deneme koşullarında gerçek ilişki düzeylerine en yakın sonuçları Spearman rank ilişki katsayısı vermiştir.

Pearson korelasyonu, Spearman rank korelasyonundan her zaman biraz daha büyüktür. Populasyondaki gerçek ilişki katsayısı arttıkça birbirine yakın sonuçlar veren bu katsayılar Kendall tau-b, Kendall tau-c ve Somer'in d katsayılarından daha iyi sonuçlar vermiştir.

Tüm bulgular genelleştirildiğinde hem tablo boyutunun hem de örneklem genişliğinin ilişki katsayıları üzerinde etkili faktörler oldukları görülmektedir. Bu sebeple araştırıcıların kullanacakları ilişki katsayısını belirlerken değişken tipinin yanı sıra tablo boyutu ve örneklem genişliğini de dikkate almaları önerilir.

7. KAYNAKLAR

Astola, J. & Virtanen I. (1981). Entropy correlation coefficient, a measure of statistical dependence for categorized data. Lappeenranta University of Technology, Department of Mathematics and Physics, Research Report No: 4

Alpar, R. (2001). Spor bilimlerinde uygulamalı istatistik. Nobel Kitabevi, Ankara, 331 s.

Alpar, R. (2006). Spor bilimlerinde uygulamalı istatistik. Nobel Yayın Dağıtım, Ankara, 360 s.

Akgül, A. (2003). Korelasyon analizi. Tıbbi araştırmalarda istatistiksel analiz teknikleri SPSS uygulamaları. Emek Ofset Ltd. Şti, Ankara, 602 s.

Altunışık, R., Coşkun R., Bayraktaroğlu S. & Yıldırım, E. (2005). Sosyal bilimlerde araştırma yöntemleri. Sakarya Kitabevi, Sakarya, 365 s.

Arıcıgil Çilan, Ç. (2009). Sosyal bilimlerde kategorik verilerle ilişki analizi. Pegem Akademi Yayınevi, Ankara, 216 s.

Altaş, D., Kaspar, EÇ. & Ergüt, Ö. (2012). İlişki katsayılarının karşılaştırılması: bir simülasyon çalışması. Namık Kemal Üniversitesi, Sosyal Bilimler Metinleri, *Süreli Hakemli Dergi*, 5, 1-10.

Anonim, (2015). Korelasyon (corralation). Alfa istatistik, İstanbul. https://www.slideshare.net/alfaistatistik/korelasyon-analizi (Erişim tarihi: (23.12.2019).

Anonim, (2019a). Contingency tables (crosstabs / chisquare test). NCSS Statistical Software. https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Contingency_Tables-Crosstabs-Chi-Square_Test.pdf (Erişim tarihi: 29.12.2019).

Anonim, (2019b). Conduct and interpret a point biserial correlation. Statistics Solutions, Florida, ABD. https://www.statisticssolutions.com/point-biserial-correlation/ (Erişim tarihi: 09.01.2020).

Anonim, (2020). Belirsizlik katsayısı. https://tr.qwe.wiki/wiki/Uncertainty_coefficient (Erişim tarihi: 19.06.2020).

Aslan, S., Akyol, FN., Dibooğlu, S., Kantarcı, B. & Serim Korkmaz, D. (2019). Sağlık alanında kullanılan ilişki katsayıları. https://docplayer.biz.tr/350674-Saglik-alaninda-kullanilan-iliski-katsayilari.html (Erişim tarihi: 17.12.2019).

Borg, W. & Gall, MD. (1989). Educational research an introduction. 5th edition. Newyork and London: Longman, 590 pp.

Bruning, JL. & Kintz, BL. (1993). İstatistik. (Çev: A. Dönmez), Gündoğan Yayınları, Ankara, 365 s.

Balcı, A. (1995). Sosyal bilimlerde araştırma yöntem, teknik ve ilkeler. Pegem Akademi, Ankara, 387 s.

Büyüköztürk, Ş., Çokluk, Ö. & Köklü, N. (2000). Sosyal bilimler için istatistik. Pegem Akademi, Ankara, 260 s.

Bryman, A. & Cramer, D. (2005). Quantitative data analysis with spss 12 and 13: a guide for social scientists. London, Newyork, 367 pp.

Borg, G. (2006). Quantitative and qualitative methods of research in educational and psychological siences. (Çev: Ahmad Reza Nasr ana others) Shahid Beheshti University and (samt) publication.

Blaikie, N. (2013). Analyzing quantitative data. Sage Publications Inc, Califonia, ABD, 352 pp.

Daniel, WW. (1990). Applied nonparametric statistics, pws-kent puplishing company, Boston.

Dawson, B. & Trapp, RG. (2004). Basic and clinical biostatistics. Lange Medical Books/McGraw-Hill, Third Edition, 438 pp.

De Muth, JE. (2006). Basic statistics and pharmaceutical statistical aplications. Second Edition. London, Newyork, 721 pp.

Diker, AC. (2009). Pearson korelasyon katsayısının tahmin edicilerinin ve bu tahmin edicilere dayanan test istatistiklerinin karşılaştırılması, Yüksek Lisans Tezi, Ege Üniversitesi, Fen Bilimleri Enstitüsü, İzmir.

Demür, E. (2014). İstatistiksel korelasyon kavramı. Yüksek Lisans Tezi, Ordu Üniversitesi, Fen Bilimleri Enstitüsü, Matematik Anabilim Dalı, Ordu.

Dikmen, AC. (2016). Ordinal ölçekli ilişki katsayıları ile kpss eğitimi veren özel eğitim kurumlarındaki öğrencilerin memnuniyetinin araştırılması sivas ili örneği. Yüksek Lisans Tezi, Cumhuriyet Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Ana Bilim Dalı, Sayısal Yöntemler Bilim Dalı, Sivas.

Ergün, M. (1995). Bilimsel Araştırmalarda bilgisayarla istatistik uygulamaları spss for windows. Ocak Yayınları, Ankara, 292 s.

Everitt, B. (2001). Statistics for psychologists. Lawrance Erlbaum Associates, Inc., New Jersey, 378 s.

Fowler, R. (1987). Power and robustness in product-moment correlation. *Applied Psychological Measurement* 11(4), 419-428.

Field, A. (2000). Discovering statistics using spss for windows. Sage Publications, London, 495 pp.

Goodman, LA. & Kruskal, WH. (1972). Measures of association for cross classifications, IV: simplification of asymptotic variances. *Journal of the American Statistical Association*, 67, 410-420.

Gamgam, H. & Altunkaynak, B. (1989). Parametrik olmayan yöntemler. Seçkin Yayıncılık, Ankara, 464 s.

Gwet, K. (2002). Kappa statistics is not satisfactory for assessing the extent of agreement between raters. Series: Statistical Methods Inter-Rater Reliability Assessment, 1, 1-5.

Gwet, KL. (2008). Computing inter-rater reliability and it’s variance in the presence of high agreement. *Br J Mathem Stat Psychol,* 61, 29-48.

Gözükara Bağ, HG., Karabulut, E. & Alpar, CR. (2010). 2×2 Tablolarda gözlemciler/gözlemler arası uyumun değerlendirilmesi. *Hacettepe Diş Hekimliği Fakültesi Dergisi,* 34 (1-2), 46-52.

Göktaş, A. & İşçi, Ö. (2011). A comparison of the most commonly used measures of association for doubly ordered square contingency tables via simulation. *Metodoloski zvezki,* 8(1), 17-37.

Hinton, PR., Brownlow, C., McMurray, I. & Cozens, B. (2004). SPSS explained. London, Newyork, 316 pp.

Hamarat, B. (2017). SPSS ve MINITAB uygulamalı istatistik. Paradigma Akademi Yayın Dağıtım, Çanakkale.

İsrael, D. (2008). Data analysis in business research a step-by-step nonparametric approach. Sage Publications Inc, India, 281 pp.

Kendall, MG. (1938). A new measure of rank correlation. *Biyometrika,* 30(1-2), 81-93.

Kundel, HL. & Polansky, M. (2003). Measurement of observer agreement. *Statistical Concepts Series.* 228, 303-308.

Kartal, M. (2006). Bilimsel araştırmalarda hipotez testleri, Nobel Yayın Dağıtım, Ankara.

Kaşko, Y. (2007). Çoklu bağlantı durumunda ikili (binary) lojistik regresyon modelinde gerçekleşen I.tip hata ve testin gücü. Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

Kesici, T. & Kocabaş, Z. (2007). Biyoistatistik. Ankara Üniversitesi, Eczacılık Fakültesi, Ders Kitabı, Ankara, 369 s.

Kılıç, S. (2009). Ölçümlerin uyumluluğu ve tıptaki uygulamaları. Yüksek Lisans Tezi, Çukurova Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı, Ankara.

Kalaycı, Ş., Demirgil, H., Sungur, O., Albayrak, AS., Eroğlu, A., Küçüksille, E., Ak, B., Karaatlı, M., Keskin, HÜ., Çiçek, EU., Kayış, A., Antalyalı, ÖL., Uçar, N. & İşler, DB. (2010). SPSS uygulamalı çok değişkenli istatistik teknikleri. Asil Yayın Dağıtım, Ankara, 426 s.

Karagöz, Y. (2010a). İlişki katsayıları ile öğrenci başarısını etkileyen faktörlerin belirlenmesi. Elektronik Sosyal Bilimler Dergisi, 9(32): 425-446.

Karagöz, Y. (2010b). SPSS 18'deki cross tabulation menüsünde geçen ilişki katsayıları. Detay Yayıncılık, Ankara, 122 s.

Karagöz, Y. (2010c). Nonparametrik tekniklerin güç ve etkinlikleri. *Elektronik Sosyal Bilimler Dergisi,* 9(33): 18-40.

Kanik, EA., Erdoğan, S. & Orekeci Temel, G. (2012). İki sonuçlu tanı testlerinde iki hekim arasındaki uyum istatistiklerinin prevalanstan etkilenme durumları. *İnönü Üniversitesi Tıp Fakültesi Dergisi,* 19 (3), 153-158.

Kocabaş, Z., Özkan, MM. & Başpınar, E. (2013). Temel biyometri. Ankara Üniversitesi, Ziraat Fakültesi, Ders Kitabı, Ankara, 381 s.

Kılıç, S. (2015). Kappa testi. *Journal of Mood Disorders,* 5 (3), 142-144.

Kılıç Topal, K. (2015). Genetik Algoritmaya dayalı yeni bir sağlam korelasyon katsayısı. Yüksek Lisans Tezi, Ondokuz Mayıs Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Samsun.

Koçyiğit, S. (2016). Gelir yaşam koşullarının nominal ilişki ölçekleri ile belirlenmesi: sivas örneği. Yüksek Lisans Tezi, Cumhuriyet Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Ana Bilim Dalı, Sivas.

Knutsen, O. (2017). Social structure, value orientations and party choice in western europe. Oslo Üniversity, Oslo, Norveç, 302 pp.

Köse, N. (2020). Araştırma yöntemleri. İstanbul Gelişim Üniversitesi, Ekonomi ve Finans Bölümü. https://docplayer.biz.tr/111646596-Arastirma-yontemleri-prof-dr-nezir-kose-istanbul-gelisim-universitesi-ekonomi-ve-finans-bolumu.html (Erişim tarihi: 22.06.2020).

Landis, JR. & Koch, GG. (1977). The measurement of observer agreement for categorical data. *Biometrics,* 33, 159–174.

Levin, J. & Fox, JA. (1991). Elementary statistics. HarperCollins Publishers Inc, New York, 501 pp.

Lebe, F. (2006). Tüketici davranış ve tercihlerinin analizi: erzurum için bir uygulama. Yüksek Lisans Tezi, Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü, İktisat Anabilim Dalı, Erzurum.

Mueller, JH., Schuessler, KF. & Costner, HL. (1970). Statistical reasoning in sociology. Boston: Houghton Mifflin.

Machin, D., Campbell, MJ., Tan, SB. & Tan, SH. (2009). Sample size tables for clinical studies. Wiley-Blackwell, Singapore, 264 pp.

Mendeş, M. (2012). Uygulamalı bilimler için istatistik ve araştırma yöntemleri. Kriter Yayınevi, İstanbul, 629 s.

Nie, NH., Hull, CH., Jenkins, JG., Steinbrenner, K. & Bent, DH. (1975). Statistical package for the social sciences: SPSS. New York: McGraw-Hill, 675 pp.

Nehmzow, U. (2006). Scientific methods in mobile robotics: Quantitative analysis of agent behaviour. Springer Science and Busines Media, 210 pp.

Nakip, M. (2013). Pazarlama araştırmaları teknikler ve SPSS destekli uygulamalar. Seçkin Yayıncılık, Ankara, 696 s.

Oktay, E. (2003). Kontenjans tablolarından elde edilen ilişki ölçüleri, Aktif Yayınevi, İstanbul, 167 s.

Öztuna, D., Elhan, AH., Kurşun, N. (2008). Sağlık araştırmalarında kullanılan ilişki katsayıları, *Turkiye Klinikleri J Med Sci,* 28(2), 160-165.

Powers, DA. & Xie, Y. (2008). Statistical methods for categorical data analysis. Emarald Group Publishing Limited, United Kingdom, North America, Japan India, Malaysia, China, 317 pp.

Puth, MT., Neuhauser, M. & Ruxton, GD. (2015). Effective use of Spearman's and Kendall's correlation coefficients for association between two measured traits. *Animal Behaviour* 102, 77-84.

Stuart, A. (1953). The estimation and comparison of association in contingency tables. *Biometrika,* 40, 105-110.

Somers, RH. (1962). A new asymmetric of association for ordinal variables. *American Sociological Review,* 27, 804.

Siegel, S. & Castellan, NJ. (1988). Nonparametric statistics for the behavioral sciences. 2nd ed. New York: McGraw-Hill, 399 pp.

Sim, J. & Wright, CC. (2005). The kappa statistic in reliability studies: use, ınterpretation, and sample size requirements. *Physical Therapy*, 85 (3), 257-268.

Sen, YS. (2012). İstatistik terimler sözlüğü. http://istatistiknotlarim.blogspot.com/2012/04/istatistik-terimler-sozlugu.html (Erişim tarihi: 14.01.2020).

Shaldehi, AH. (2013). Using Eta (η) correlation ratio in analyzing strongly nonlinear relationship between two variables in practical researches. *Journal of mathematics and computer science,* 7, 213-220.

Sümbüloğlu, V., Alpar, R. & Özdemir, P. (1998). Değişkenler arası ilişkilerin incelenmesi. Hacettepe Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Ankara, http://www.ichastaliklaridergisi.org/managete/fu_folder/1998-06/html/1998-5-6-416-419.html (Erişim tarihi: 23.06.2020).

Şencan, H. (2005). Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik. Seçkin Yayıncılık, Ankara, 867 s.

Şencan, H. (2007). Sosyal ve davranışsal bilimlerde bilimsel araştırma. Seçkin Yayıncılık, Ankara, 170 s.

Taştan, SN. (2013). Kategorik-sınıflandırılmış veri analiz yöntemleri ve uygulamaları. Yüksek Lisans Tezi, Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı, Isparta.

Tuğran, E. (2015). Bazı korelasyon katsayılarının I.Tip hata ve testin gücü bakımından karşılaştırılması. Yüksek Lisans Tezi, Çanakkale Onsekiz Mart Üniversitesi, Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı, Çanakkale.

Temel, G. & Erdoğan, S. (2017). Determining the sample size in agreement studies. *Marmara Medical Journal*, 30, 100-112.

Terzi, Y. (2019). SPSS ile istatistiksel veri analizi. Samsun, https://personel.omu.edu.tr/docs/ders_dokumanlari/1029_93417_1500.pdf (Erişim tarihi: 16.12.2019).

Üçkardeş, F. (2006). İstatistik testler üzerine bir çalışma. Yüksek Lisans Tezi, Kahramanmaraş Sütçü İmam Üniversitesi, Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı, Kahramanmaraş.

Üstün, B. (2016). Araştırma tipleri. http://www.phdernegi.org/wp-content/uploads/2016/03/araştırma_türleri.pdf (Erişim tarihi: 23.06.2020).

Yenigün, DC. (2007). A test of ındependence in two-way contingency tables based on maximal correlation. Doctoral Thesis, Bowling Green State Univesity, Bowling Green Institute, Philosophy, Ohio/ABD.

ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı	Sinem ŞENSOY
Doğum Yeri	ORDU
Doğum Tarihi	09.04.1990
Uyruğu	☒ T.C. ☐ Diğer:
Telefon	0542 340 18 42
E-Posta Adresi	sinem.sensoy@yahoo.com
	sinemsy@outlook.com.tr

Eğitim Bilgileri	
Önlisans	
Üniversite	Marmara Üniversitesi
Yüksekokul	Teknik Bilimler Meslek Yüksekokulu
Bölümü	Su Ürünleri
Mezuniyet Yılı	05.07.2011
Lisans	
Üniversite	Atatürk Üniversitesi
Fakülte	Ziraat Fakültesi
Bölümü	Zootekni
Mezuniyet Yılı	03.06.2016