

ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

76654

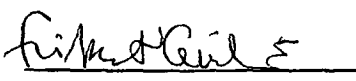
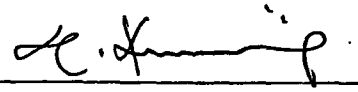
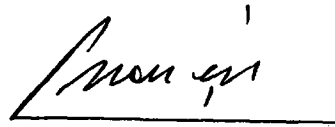
VARYANS ANALİZİNİN ÖNŞARTLARI VE TRANSFORMASYONLAR

Benian TEKİNDAL

DOKTORA TEZİ
ZOOTEKNİ ANABİLİM DALI

76654

Bu tez 02/03/1998 tarihinde aşağıdaki jüri tarafından
edilerek Oybirliği/Oyçokluğu ile kabul edilmiştir.

		
Prof. Dr. Fikret GÜRBÜZ (Danışman)	Prof. Dr. Tahsin KESİCİ	Doç. Dr. Burhan CİL

**VARYANS ANALİZİNİN ÖNŞARTLARI VE
TRANSFORMASYONLAR**

Benian TEKİNDAL

**DOKTORA TEZİ
ZOOTEKNİ ANABİLİM DALI**

**ANKARA
1998**

ÖZET

Doktora Tezi

VARYANS ANALİZİNİN ÖNŞARTLARI VE TRANSFORMASYONLAR

Benian TEKİNDAL

**Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Zootekni Anabilim Dalı**

**Danışman: Prof.Dr.Fikret GÜRBÜZ
1998, Sayfa: 70**

**Jüri: Prof.Dr.Fikret GÜRBÜZ
Prof.Dr.Tahsin KESİCİ
Doç.Dr.Burhan ÇİL**

Bu tez çalışmasında, varyans analizi tekniğinin önşartlarının sağlanmadığı koşullara ilişkin ve ilgili önşartları sağlayacak transformasyonlar üzerinde ayrıntılı açıklamalar yapılmıştır.

Ayrıca simülasyonla elde edilen normal dağılım gösteren verilerde varyans analizi tekniğinin önşartlarından biri olan varyansların homojenlik önşartı yerine gelmediği durumlarda, varyans analizi tekniği uygulanmış ve gerçekleşen 1. tip hata olasılıkları tespit edilmiştir. Daha sonra bu veriler logaritmik, karekök ve ters transformasyonlara tabi tutularak varyans analizi tekniği uygulanmış ve 1. tip hata oranlarında meydana gelen değişmelere bakılmıştır.

Eşit ve farklı gözlem sayıları içeren 2, 3 ve 4 grupta, değişik varyans oranları ele alınmış ve sonuçta yukarıda bahsedilen 3 transformasyonun, 1. tip hatada istenilen doğrultuda değişiklik yapmadığı gözlenmiştir.

χ^2 dağılımı gösteren verilerde 1, 3, 5 ve 15 serbestlik derecesi dikkate alınarak üretilen veriler eşit ve farklı sayılarda gözlem içeren 2, 3, 4 grup halinde düzenlendikten sonra hem orjinal verilere hem de transforme edilmiş olanlara

varyans analizi tekniđi uygulanmış ve sonuçta dağılımın şeklinin eğri olduđu bu koşullarda uygulanan transformasyonlardan karekök ve logaritmik transformasyonların 1. tip hata olasılığının başlangıçta belirlenen % 5 seviyesinde kalmasını sağladığı tespit edilmiştir.

Sonuç olarak logaritmik, karekök ve ters transformasyonların, varyansların heterojen olması durumunda etkili bulunmadığı dağılımın şeklinin eğik olması durumunda yani χ^2 dağılımında ise, logaritmik ve karekök transformasyonlarının 1. tip hata olasılıklarını kararlaştırılan düzeye yaklaştırmak bakımından etkili sonuçlar verdiği saptanmıştır.

Çalışmada her bir dağılım şekli, grup sayısı, örnek genişliği ve transformasyon tipi kombinasyonu için hesaplamalar üretilen 100.000'er veri setinde gerçekleştirilmiştir.

ANAHTAR KELİMELELER: Normal dağılım, χ^2 dağılım, transformasyonlar, varyans analizi tekniđi, önşartlar, Box-cox tekniđi

ABSTRACT

Ph.D. Thesis

**ASSUMPTIONS OF ANALYSIS OF VARIANCE AND
TRANSFORMATIONS**

Benian TEKİNDAL

**Ankara University
Graduate School of Natural and Applied Sciences
Department of Animal Science**

**Supervisor: Prof.Dr.Fikret GÜRBÜZ
1998, Page: 70**

**Jury: Prof.Dr.Fikret GÜRBÜZ
Prof.Dr.Tahsin KESİCİ
Assoc.Prof.Dr.Burhan ÇİL**

A detailed explanations were made on ANOVA with its assumptions satisfied and transformations under unsatisfied conditions.

Analysis of variance was made on normally distributed simulated data with heterogenous variance and the probability of type I error was determined. The same type of error was also observed in the same data with logarithmic, square-root and inverse transformations. Analysis applied to 2, 3 and 4 groups including equal and unequal observations showed that above mentioned types of transformations did not result in any change in Type I error.

ANOVA was also applied to 2, 3 and 4 groups of equal and unequal observations simulated to have 1, 3, 5 and 15 degrees of freedom with χ^2 distrubution and it was found that the square-root and logarithmic transformations kept Type I error at predetermined 5 % level.

As a result; logarithmic square-root and inverse transformations were not found effective in heterogenous variances in normal distribution wehere as logarithmic and square-root transformations were effective in χ^2 distributions.

This study was conducted in 100.000 data set.

Key Words: Normal distribution, chi-square distribution, transformations, analysis of variance technique, assumptionsBox-cox technique



ÖNSÖZ VE TEŞEKKÜR

Uygulamalı bilimlerde yaygın olarak kullanılan varyans analizi tekniğinin uygulanabilmesi için belirli önşartların yerine gelmiş olması gerekir. Bu çalışmada varyans analizi tekniğinin ön şartlarının kapsamlı bir şekilde incelenmesi ve bu ön şartlardan en önemlisi olan "Varyansların Homojenliği" ön şartından olan farklı derecelerdeki ayrılışların, karşılaştırılacak grup sayısı ve gruplardaki eşit olmayan gözlem sayılarını da dikkate alarak, gerçekleştirecek I. tip hata olasılıklarının irdelenmesi ve varyansların homojenliği ön şartını gerçekleştirdiği kabul edilen transformasyon tekniklerinin ele alındığı bu çalışmanın yeni araştırmalar için yararlı olacağını ümit ediyorum.

Bana bu konuda çalışma olanağı veren ve çalışmalarımnda her zaman değerli fikirleriyle bana yol gösteren danışmanım, Sayın Prof.Dr.Fikret GÜRBÜZ'e ve Araştırma Görevlisi Sayın Arzu OKUL'a, Araştırma Görevlisi Sayın Handan ÇAMDEVİREN'e ve tezimin yazımını gerçekleştiren Uzman Mücahit ŞAR'a teşekkürlerimi sunarım.

İÇİNDEKİLER

ÖZET	i
ABSTRACT	iii
ÖNSÖZ VE TEŞEKKÜR	v
ŞEKİLLER DİZİNİ	ix
ÇİZELGELER DİZİNİ	x
1. GİRİŞ	1
2. KURAMSAL TEMELLER VE KAYNAK ARAŞTIRMASI	3
2.1. Varyans Analizi	3
2.1.1. Normallik ön şartı	8
2.1.2. Etkilerin eklenebilirliği ön şartı	10
2.1.3. Bağımsızlık ön şartı	10
2.1.4. Varyansların homojenliği ön şartı	11
2.1.5. Varyans analizi tekniğinin ön şartları gerçekleşmediği taktirde karşılaşılan durumlar	12
2.2. Transformasyonlar	13
2.2.1. Güç transformasyonları	16
2.2.2. Üstel transformasyon	19
2.2.3. Ters transformasyonlar	21
2.2.4. Karekök transformasyon	22
2.2.5. Logaritmik transformasyon	22
2.2.6. Arcsin transformasyon	23
2.2.7. Hiperbolik transformasyon	23
2.2.8. Açısal transformasyon	24
2.2.9. İhtimal integral transformasyonu	24
2.3. Kaynak Araştırması	28
2.3.1. Genel açıklamalar	28

3. MATERYAL VE METOD	34
3.1. Materyal	34
3.2. Metod	34
4. SONUÇLAR VE TARTIŞMA	38
4.1. Normal Dağılım Gösteren Veriler	38
4.1.1. İki grup ve gruplardaki gözlem sayıları eşit iken elde edilen sonuçlar	38
4.1.2. İki grup ve gruplardaki gözlem sayıları birbirinden farklı iken elde edilen sonuçları	41
4.1.3. Üç grup ve gruplardaki gözlem sayıları birbirine eşit iken elde edilen sonuçlar	43
4.1.4. Üç grup ve gruplardaki gözlemler birbirinden farklı iken elde edilen sonuçlar	46
4.1.5. Dört grup ve gruplardaki gözlem sayıları birbirine eşit iken elde edilen sonuçlar	47
4.1.6. Dört grup ve gruplardaki gözlem sayıları birbirinden farklı iken elde edilen sonuçlar	50
4.2. χ^2 Dağılımı Gösteren Veriler	52
4.2.1. Bir serbestlik dereceli χ^2 dağılımından herbirinde eşit sayıda gözlem bulunan gruplara ilişkin sonuçlar	52
4.2.2. Bir serbestlik dereceli χ^2 dağılımında herbirinde farklı sayıda gözlem bulunan gruplara ilişkin sonuçlar	54
4.2.3. Üç serbestlik dereceli χ^2 dağılımında, herbirinde eşit sayıda gözlem bulunan gruplarda ilişkin sonuçlar	56
4.2.4. Üç serbestlik dereceli χ^2 dağılımında, herbirinde farklı sayıda gözlem bulunan gruplara ilişkin elde edilen sonuçlar	57
4.2.5. Beş serbestlik dereceli χ^2 dağılımında, herbirinde eşit sayıda gözlem bulunan gruplara ilişkin elde edilen sonuçlar	58
4.2.6. Beş serbestlik dereceli χ^2 dağılımında, birbirinde farklı sayıda gözlem bulunan gruplara ilişkin sonuçlar	60

4.2.7. Onbeş serbestlik dereceli χ^2 dağılımında, herbirinde eşit sayıda gözlem bulunan gruplara ilişkin sonuçlar	61
4.2.8. Onbeş serbestlik dereceli χ^2 dağılımında, herbirinde farklı sayıda gözlem bulunan gruplara ilişkin sonuçlar	62
4.3. Genel Değerlendirmeler	64
KAYNAKLAR	67



ŞEKİLLER DİZİNİ

Şekil 4.1. Grup sayısı 2 ve gruplardaki gözlem sayıları eşit olduğunda 100000 deneme sonucunda gerçekleşen I. tip hata ihtimalleri	40
Şekil 4.2. Grup sayısı 2 ve gruplardaki gözlem sayıları birbirinden farklı olduğundan 100000 deneme sonucunda red edilen hipotez sayılarının dağılımı	43
Şekil 4.3. Grup sayısı 3 ve gruplardaki gözlem sayıları eşit olduğunda 100000 deneme sonucunda red edilen hipotez sayılarının dağılımı	45
Şekil 4.4. Üç grup gruplardaki gözlemler birbirinden farklı olduğunda 100000 deneme sonucunda red edilen hipotez sayılarının dağılımı	47
Şekil 4.5. Grup sayısı 4 ve gruplardaki gözlem sayıları eşit olduğundan 100000 deneme sonucunda red edilen hipotez sayılarının dağılımı	49
Şekil 4.6. Grup sayısı 4 ve gruplardaki gözlem sayıları birbirinden farklı olduğunda 100000 deneme sonucunda red edilen hipotez sayılarının dağılımı	51

ÇİZELGELER DİZİNİ

Çizelge 2.1.	Transformasyonların özeti	27
Çizelge 3.1.	Normal dağılımlı populasyonlara ilişkin simulasyon şeması	37
Çizelge 3.2.	χ^2 dağılımlı populasyonlara ilişkin simulasyon düzenleri	37
Çizelge 4.1.	Grup sayısı 2 ve gruplardaki gözlem sayıları eşit olduğunda 100000 deneme sonucunda red edilen hipotez sayıları	39
Çizelge 4.2.	Grup sayısı 2 ve gruplardaki gözlem sayıları birbirinden farklı olduğunda 100.000 deneme sonucunda red edilen hipotez sayıları	41
Çizelge 4.3.	Grup sayısı 3 ve gruplardaki gözlem sayıları eşit olduğunda 100.000 deneme sonucunda red edilen hipotez sayıları	44
Çizelge 4.4.	Grup sayısı 3 ve gruplardaki gözlem sayıları birbirinden farklı olduğunda 100.000 deneme sonucunda red edilen hipotez sayıları	46
Çizelge 4.5.	Grup sayısı 4 ve gruplardaki gözlem sayıları eşit olduğunda 100.000 deneme sonucunda red edilen hipotez sayıları	48
Çizelge 4.6.	Grup sayısı 4 ve gruplardaki gözlem sayıları birbirinden farklı olduğunda 100.000 deneme sonucunda red edilen hipotez sayıları	50
Çizelge 4.7.	Bir serbestlik dereceli χ^2 dağılım gösteren veriler ve grup sayısı 2 ve gruplardaki gözlem sayısı eşit olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayıları	52
Çizelge 4.8.	Üç serbestlik dereceli χ^2 dağılımı gösteren 2, 3 ve 4 grupta ve gruplardaki gözlem sayılarının farklı iken 100.000 deneme sonucunda red edilen hipotez sayıları	54
Çizelge 4.9.	Üç serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı eşit olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayıları	56

- Çizelge 4.10. Üç serbestlik dereceli χ^2 dağılımı gösteren veriler
ve gruptaki gözlem sayısı farkı olduğu durumda 100.000
deneme sonucunda red edilen hipotez sayıları 58
- Çizelge 4.11. Beş serbestlik dereceli χ^2 dağılımı gösteren veriler ve
gruptaki gözlem sayısı eşit olduğu durumda 100.000 deneme
sonucunda red edilen hipotez sayısı 59
- Çizelge 4.12. Beş serbestlik dereceli χ^2 dağılımı gösteren veriler ve
gruptaki gözlem sayısı farklı olduğu durumda 100.000 deneme
sonucunda red edilen hipotez sayısı 60
- Çizelge 4.13. Onbeş serbestlik dereceli χ^2 dağılımı gösteren veriler
ve gruptaki gözlem sayısı eşit olduğu durumda 100.000
deneme sonucunda red edilen hipotez sayısı 61
- Çizelge 4.14. Onbeş serbestlik dereceli χ^2 dağılımı gösteren veriler
ve gruptaki gözlem sayısı farklı olduğu durumda 100.000
deneme sonucunda red edilen hipotez sayısı 63

1. GİRİŞ

Transformasyonun kelime anlamı dönüşümdür. Yani yapıyı değiştirme ve düzenleme anlamına gelmektedir. Veri transformasyonu ise mevcut verilerin uygun şekillerde değiştirilerek yeni bir yapı kazanmasıdır. Bu haliyle transformasyon, verilerin istatistiksel metodlarla değerlendirilmesinde önemli bir yere sahiptir. Aslında transformasyon istatistiksel bir hile gibi düşünülse bile gerçekte durum böyle değildir. Çünkü, uygulanacak transformasyon bir grup veriye değil analize tabi tutulacak bütün verilere uygulanmakta ve verilerin hepsini etkilemektedir.

Çalışmalarda, elde edilen verilere uygulanan muamelenin etkisini araştırmak amacıyla kullanılan istatistik analiz tekniklerinden birisi varyans analiz tekniğidir. Birçok istatistik test tekniğinde olduğu gibi varyans analizi tekniğinin uygulanabilmesi için de verilerde bir takım özellikler aranır. Bu özellikler uygulanacak test tekniğinin önşartları olarak adlandırılır.

Varyans analizi tekniğinin uygulanabilmesi için belli başlı dört tane temel önşart vardır. Bu önşartlar kısaca; verilerin normal dağılım göstermesi, farklı muamele guruplarında yer alan verilerin varyanslarının homojen olması, farklı muamele guruplarında yer alan deney ünitelerinin birbirinden farklı olması (bağımsızlık) ve iki ve daha fazla faktör içeren denemelerde faktör etkilerinin eklemeli olması şeklinde tanımlanır. Bu önşartları gerçekleştiren verilere varyans analizi tekniği uygulanabilir. Bir veya daha fazla önşartın gerçekleşmediği durumlarda ise verilere ya başka istatistik analiz tekniği uygulanır veya uygun bir test tekniği düşünülemediği durumda veriler değerlendirilemez. En uygun analiz tekniği belirlerken göz önüne alınan nokta ise başlangıçta belirlenmiş olan 1. tip hata ihtimalini (α) değiştirmeyecek analiz tekniğidir. Yanlış bir analiz tekniği ile verileri değerlendirmek, hiç değerlendirmemekten daha kötü sonuçlar verir araştırmacıyı yanlış yorumlamalara itmiş olur.

Sayılan bu durumlar dikkate alındığında doğru seçim yapmak verileri deęerlendiren ve deęerlendirme sonularını yorumlayan kiřilere kalmıřtır.

Yukardaki sorulara en doęru cevabı verecek yntemi seerken gerek arařtırıcının gerekse verileri deęerlendiren kiřinin gz nne alması gerekli olan ok nemli hususlar vardır. Mesela, testin kolay uygulanabilirlięi, analiz sonucunda elde edilen sonuların yorumlanabilirlięinin kolay olması, yorumlanan sonulardaki hata olasılıęının bařlangıta belirlenen seviyeyi ařmaması, uygulanan testin gcnn yksek olması ve sonuların genellenebilmesi gibi ana etkilerin yanı sıra maliyetinin uygun olması, fazla zaman alıcı olmaması gibi etkilerin de gz nne alınması gerekir.

Bu tez alıřmasında, varyans analizi teknięinin n řartları gerekleřmedięi durumlarda veri yapısındaki deęiřmelerin, yani verilere uygulanacak transformasyonların varyans analizi teknięinin n řartlarını ve analiz sonucunda elde edilecek yorumları nasıl etkiledięi arařtırılmıř ve yapılan simlasyon alıřmasında varyansların homojenlik nřartı yerine gelmedięi zaman verilerin varyans analizi teknięi ile deęerlendirildięinde bunun gerekleřen 1. tip hataya etkisi arařtırılmıřtır.

2. KURAMSAL TEMELLER VE KAYNAK ARAŞTIRMASI

2.1. Varyans Analizi

R.A. Fisher tarafından 1920'lerde geliştirilen varyans analizi, bir bağımlı değişken üzerinde etkide bulunan bağımsız değişkenlerin tesirlerini ortaya koymak amacıyla oluşturulan modellerde kullanılan bir istatistik tekniğidir.

Varyans analizinde bağımsız değişkene "Faktör" ve bağımsız değişkenin aldığı değerlerde "Faktör düzeyleri" denir. Misal olarak; cinsiyet bir faktördür. Faktörün düzeyleri ise erkek ve dişidir. Aynı şekilde "Doğu Anadolu Bölgesi" faktör, bu bölgedeki şehirler (Malatya, Erzurum, Elazığ, Kars) faktörün düzeyleri olabilir.

Varyans analizinin uygulandığı modellerde bir veya daha fazla faktör olabilir ve amaç faktörlerin çeşitli düzeylerinin bağımlı değişken üzerindeki etki derecelerini ortaya çıkarmaktır. Bunun için faktörlerin uygulanacağı popülasyondan tesadüfi olarak seçilmiş bulunan bireyler gruplara ayrılır ve grupların her birine faktörlerin çeşitli seviyelerinin olabilir kombinasyonlarından biri uygulanır. Mesela, iki faktörlü bir modelde birinci faktörün r , ikinci faktörün c tane seviyesi varsa örnekteki bireyler $c \times r$ guruba ayrılabilir ve her guruba ayrı bir faktör seviye kombinasyonu tatbik olunabilir.

Gruplarda eşit sayıda deney ünitesi (gözlem) bulunması zorunlu olmamasına karşın eşit sayıda deney ünitesi bulunduğunda bazı avantajlar sağlanmış olur. Önemli olan husus deney ünitelerinin gruplara ayrılmasında tesadüfi dağıtım (randomization) kurallarına uyma zorunluluğudur.

Materyal, mukayese edilecek muamele gruplarına öyle dağıtılmalıdır ki üzerinde durulan özellik bakımından gruplar içindeki deney ünitelerinin birbirinden farklılığı (gruplar içi varyans) grupların ortalamaları arasındaki farklılığa (gruplar arası varyans) eşit olsun, bu ise ancak materyalin gruplara rastgele

dağıtılması ile sağlanır. Faktörlerin çeşitli seviyeleri farklı guruplara uygulandıktan sonra bütün deney ünitelerinde bağımlı değişkenin değeri tesbit olunur. Eğer faktör seviyeleri etki yönünden birbirinden farklı ise gurupların ortalamalarının birbirinden farklı olması beklenir. Gerçekte guruplar arasında görülen farklılıkların bir kısmı faktör seviyelerinden bir kısımda kontrol edilmemiş değişkenlerden ve ölçme hatalarından ileri gelebilir. Varyans analizi faktörlere atfedilecek etkinin önemli olup olmadığını ortaya çıkarır.

Varyans analizi tekniğinde test edilen sıfır hipotezi gurup ortalamalarının birbirine eşit olduğu yani faktör seviyelerinin etkileri arasında önemli bir fark bulunmadığı hipotezidir. Varyans analizinde her gurup ayrı bir populyasyondan seçilmiş bir örnek olarak ele alınır. Böylece test edilen sıfır hipotezi çok sayıdaki populyasyonların ortalamalarının birbirine eşit olduğunu ifade eder.

Varyans analizinde etkileri incelenen faktörlere ait seviyeler bazen belirli sınıflandırmalardan ibarettir.

Faktör seviyeleri, faktör olarak alınan değişkenin değer aralığından keyfi olarak belirlenen değerler olarak saptanır. Böyle durumlar sabit etki (fixed effects) modelleri olarak isimlendirilir.

Eğer faktör seviyeleri, değişkenin değerleri arasından tesadüfi olarak seçilen değerler olarak belirlenirse tesadüfi etki (random effects) modelleri olarak isimlendirilir. Sabit etki modellerinde varılan sonuç sadece modelde yer verilmiş seviyeler için geçerli olduğu halde tesadüfi etki modellerinde elde edilen sonuç faktörün bütün seviyeleri için geçerlidir. İki veya daha fazla faktörün yer aldığı modellerde bazı faktörlerin seviyeleri sabit etki bazı faktörlerin seviyeleri de tesadüfi etki modeline uygun olarak belirlenmiş olduğu durumlarda bu modellere karışık etki (mixed effects) modelleri olarak isimlendirilir. Varyans analizinde, bağımlı değişkenin toplam varyansı (total variance), guruplar arası varyansı

(between groups variance), gruplar içi varyansı (within group variance), gibi kavramlardan yararlanılır (Avrallıoğlu 1977).

Toplam varyans, incelenen bütün bireyler birarada ele alındığında hesaplanan varyanstır.

Gruplar arası varyans ise grupların ortalamaları arasındaki farklılığın bir ölçüsüdür.

Gerçekte varyans analizinin toplanabilir kısımlara ayrıldığı kavram "kareler toplamı" (sum of squares) kavramıdır. Kareler toplamı, sözü edilen varyansların hesabında payda yer alan değerdir.

Varyans analizinin matematiksel yönünü Sezgin (1972) de aşağıdaki gibi ifade etmiştir.

Durumu tek faktörlü sınıflama için ele alırsak r muameleden ibaret bir denemede her muamele n defa denenmişse,

$\bar{X}_{..}$: Genel ortalama

$\bar{X}_i$: i'inci muamelenin ortalaması

X_{ij} : i'inci muameledeki j'inci gözlem olmak üzere

$$(X_{ij} - \bar{X}_{..}) = [(X_{ij} - \bar{X}_i) + (\bar{X}_i - \bar{X}_{..})] \dots\dots\dots (2.1)$$

eşitliği geçerlidir.

Her iki tarafın karesi alınırsa;

$$(X_{ij} - \bar{X}_{..})^2 = (X_{ij} - \bar{X}_i)^2 + 2(X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X}_{..}) + (\bar{X}_i - \bar{X}_{..})^2 \dots\dots\dots (2.2)$$

çıkar. Şimdi i ve j indisleri üzerinden eşitliğin her iki yanını toplarsak;

$$\sum_{i=1}^r \sum_{j=1}^n (X_{ij} - \bar{X}_{..})^2 = \sum_{i=1}^r \sum_{j=1}^n (X_{ij} - \bar{X}_{i.})^2 + \sum_{i=1}^r (\bar{X}_{i.} - \bar{X}_{..})^2 \cdot n \quad (2.3)$$

bulunur. Çünkü;

$$\sum_{i=1}^r \sum_{j=1}^n (X_{ij} - \bar{X}_{i.})(\bar{X}_{i.} - \bar{X}_{..}) = \sum_{i=1}^r [(\bar{X}_{i.} - \bar{X}_{..}) \sum_{j=1}^n (X_{ij} - \bar{X}_{i.})]$$

$$\text{her } i \text{ değeri için } \sum_{j=1}^n (X_{ij} - \bar{X}_{i.}) = 0$$

olduğu ortalamanın tarifinden anlaşılır.

Aynı yaklaşımı kullanarak r sıra ve s sütundan meydana gelmiş iki yönlü bir sınıflamada i sıraları j sütunları göstermek üzere;

$$\sum_{i=1}^r \sum_{j=1}^s (X_{ij} - \bar{X}_{..})^2 = s \sum_{i=1}^r (\bar{X}_{i.} - \bar{X}_{..})^2 + r \sum_{j=1}^s (\bar{X}_{.j} - \bar{X}_{..})^2 + \sum_{i=1}^r \sum_{j=1}^s (\bar{X}_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2 \quad (2.4)$$

$$\sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^t (X_{ijk} - \bar{X}_{...})^2 = S + \sum_{i=1}^r (\bar{X}_{i..} - \bar{X}_{...})^2 + r + \sum_{j=1}^s (\bar{X}_{.j.} - \bar{X}_{...})^2$$

$$+ rs \sum_{k=1}^t (\bar{X}_{...k} - \bar{X}_{...})^2 + t \sum_{i=1}^r \sum_{j=1}^s (X_{ij.} - \bar{X}_{i..} - \bar{X}_{.j.} + \bar{X}_{...})^2$$

$$+ s \sum_{i=1}^r \sum_{j=1}^s (X_{i.k} - \bar{X}_{i..} - \bar{X}_{...k} + \bar{X}_{...})^2$$

$$+ r \sum_{j=1}^s \sum_{k=1}^t (\bar{X}_{.jk} - \bar{X}_{.j.} - \bar{X}_{...k} + \bar{X}_{...})^2$$

$$+ \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^t (X_{ijk} - \bar{X}_{ij.} - \bar{X}_{i.k} + \bar{X}_{.jk} + \bar{X}_{i..} + \bar{X}_{.j.} + \bar{X}_{..k})^2 \quad (2.5)$$

eşitliğinin geçerli olduğu görülür. Denemenin faktör sayısının artışına paralel olarak sağ taraftaki ifade de daha karmaşık durumlar arz etmektedir.

Buraya kadar çıkarılan (2.3), (2.4) ve (2.5) eşitlikleri varyans analizi işleminde elde edilen kareler toplamından başka bir şey değildir. Uygun serbestlik dereceleri ile bölünerek kareler ortalamalarını da elde etmek mümkündür.

Buraya kadar yapılan açıklamalar matematiksel olarak doğru olmakla beraber gözlemler ve ortalamaları hakkında bazı ön şartları olan modellere dayandırılarak test edilmektedir. Elde edilen sonuçların güvenilir olabilmesi için modellerin dayandıkları ön şartların sağlanması gerekmektedir.

Fisher'in varyans analizini geliştirirken göz önüne aldığı ön şartları Eisenhart (1947) de topluca özetlemiştir.

Sokol and Rohlf (1969) da Sezgin (1972), Düzgüneş vd (1985) de aynı konuya değinmişlerdir.

Canavos and Miller (1993) gerekli ön şartlara dikkat çekmişlerdir.

Bu görüşlerin çerçevesinde varyans analizi tekniğinin ön şartları aşağıdaki gibi özetlenebilir;

- a. Normal dağılım
- b. Gözlemlerin bağımsızlığı
- c. Etkilerin eklenebilirliği
- d. Varyansların homojenliği

2.1.1. Normallik ön şartı

Varyans analizi tekniğinin en önemli ön şartlarından birisi de gözlemlerin normal dağılım göstermesidir. Tek faktörün çeşitli düzeylerinin belirli bir bağımlı değişken üzerindeki etkilerinin incelendiği model tek faktörlü model diye adlandırılır ve en basit şekli ile;

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij} \dots\dots\dots (2.6)$$

gösterilir.

Bu eşitlikte yer alan y_{ij} bağımlı değişken olarak adlandırılır ve kendi gurupları etrafında rastgele dağılım gösterir. μ populasyon ortalamasını α_i parametreleri, ϵ_{ij} ise tesadüfi değişken diye adlandırılan ve hata terimini ifade eder. Hata terimi, ortalaması sıfır varyansı σ^2_ϵ olan normal dağılım gösteren bir değişkendir. Bu durumu kısaca şöyle ifade edilir.

$$\epsilon \sim N(0, \sigma^2_\epsilon)$$

Hataların normal dağılım göstermelerine ilişkin ön şart hipotez kontrolü yapılmak istendiği zaman daha ciddi inceleme yapmayı gerektirmektedir.

Normal dağılımın iki özelliği vardır. Bunlar;

1. Simetri

Bu özellikle bağımlı değişkenin ortalama etrafında tam simetrik bir şekilde yani bu değişkenlerin % 50'sinin ortalamanın sağında % 50'sinin ise ortalamanın solunda yer aldığını ifade eder. Simetriliğin ölçüsü, normal dağılımın 3. momenti olan eğiklik katsayısıdır. Bu katsayıyı γ_3 ile gösterirsek tam simetrik bir normal dağılımda bu katsayının değeri 0 dır. 0 dan sağa veya sola sapmaların ölçüsü olarak

dağılımın sağa veya sola yatık olduğundan bahsedilir, daha açık bir ifade ile $\gamma_3 > 0$, ise dağılım sağa yatık $\gamma_3 < 0$ ise dağılım sola yatıktır.

2. Basıklık veya Diklik

Normal dağılımın bu özelliği ise 4. moment olan diklik katsayısı ile ifade edilir, bu katsayıyı γ_4 ile gösterecek olursak normal dağılımda bu katsayı 3'e eşittir. $\gamma_4 > 3$ ise dağılımın dik olduğundan $\gamma_4 < 3$ ise dağılımın basık olduğundan bahsedilir.

Normallikten sapma hangi ölçüde olursa olsun örnek genişliğindeki artış kullanılan test istatistiğinde hata yapma ihtimalini azaltır, Düzgüneş vd. (1987) mesela χ^2 dağılışında örnek genişliği 30'un üzerine çıkınca söz konusu verilerin artık bir normal dağılım gösterdiği varsayılır.

Varyans analizi hipotez kontrollerine dönük olarak kullanıldığında önemlilik testleri F dağılışı yardımıyla yapılabilmektedir.

Şenyurt (1979) tarafından bildirildiğine göre; Li (1957), populasyonların normal dağılış göstermemesinin F testi veya t testinde önemli hatalara neden olamayacağını, ancak F ve t tablosu hipotez kontrollerine dönük olarak, kritik bölgelerin saptanması amacıyla kullanılacak olursa, gerçek önem seviyesinin belirtilenden bir miktar fazla olacağını ileri sürmüştür ki bu I. tip hata yapma ihtimalinin artması demektir.

Hataların dağılışının çok aşırı bir çarpıklık göstermesi halinde F testinin önemlilik seviyesinin çok fazla etkileneceği görülmüştür.

Sokol and Rohlf (1969), Eisenhart (1947) bu ön şart gerçekleşmediğinde standart normal dağılış veya F dağılışı yardımıyla doğru bir önemlilik testi yapılamayacağını belirtmiştir.

2.1.2. Etkilerin eklenebilirliği ön şartı

Uygulanabilecek her varyans analizi tekniği için bir model yazmak mümkündür, bu modelde bağımlı değişken olarak adlandırdığımız üzerinde durduğumuz özelliğe ait veriler, örneklerin çekildiği popülasyona ait ortalama etkisi araştırılan faktörler ve etkisi tesadüfi olarak kabul edilen hata unsuru yer alır. En basitinden en karmaşığına kadar bu modellerde bahsedilen bu etkiler üst üste eklenirse söz konusu modele eklemeli model denir.

2.1.3. Bağımsızlık ön şartı

Gözlemlerin bağımsızlığından anlatılmak istenen; farklı muamele guruplarında yer alan ölçüm değerlerinin farklı deney ünitelerinden elde edilmiş olmasıdır.

Gözlemlerin bağımsız olması durumunda hatalar da bağımsız olur. O zaman varyans analizi sonucunda hesaplanacak olan F istatistiğinin paydasında yer alan Hata Kareler ortalamasının beklenen değeri içerisinde hatalar arasındaki bağımlılığın göstergesi olan Kovaryans değeri 0'a eşit olur ki böylece hata varyansının yansız bir tahmin edicisi hesaplanmış olur.

Aksi durumda yani gözlemler bağımlı olduğu durumda hata varyansının beklenen değeri içerisinde kovaryans unsurunun bulunmasından dolayı hata varyansı küçülür ki bu istenmiyen bir küçülmedir ve bu da F istatistiğinin hesaplanan değerinin büyümesine yani I tip hata yapma ihtimalinin artmasına sebep olur. Yukarıdaki açıklamalara bağlı olarak şöyle bir genel sonuç çıkarmak mümkündür.

Gözlemler bağımlı iken klasik varyans analizi tekniğini kullanmak hata ihtimalini değiştirir.

2.1.4. Varyansların homojenliği ön şartı

Varyans analizi tekniğinde hesaplanacak F istatistiğinin paydasında yer alan varyans unsuru hata kareler ortalaması olarak adlandırılır ve aslında bir toplanmış (Pooled) varyanstır, bilinmeyen populasyon varyansının en iyi tahminidir. Bu varyans unsuru popülasyondan rastgele seçilen ve birbirileri ile karşılaştırılmak istenen gruplara ait varyansların ortalamalarının hesaplanması ile bulunur. Toplama işlemi ancak toplanabilen değerlere uygulanabileceğine göre, gruptaki varyanslarında toplanabilmesi için birbirlerinden toplanamayacak kadar farklı olmamaları gerekir ki buna homojenlik denir (Düzgüneş vd 1987). Hataların ortak bir varyansa sahip olmaları varsayımı ile örneği oluşturan bireylerin hata varyansına olan katkılarının büyük farklılık göstermedikleri ifade edilmektedir. Populasyondan çekilen örneklerin varyanslarının birbiri ile toplanamayacak kadar farklı olması durumunda eğer yine de bunları üst üste ekliyerek hata varyansı bulunup F istatistiği hesaplanırsa grupta ortalamaları arasındaki fark az olduğu durumda hata varyansının büyümesinden dolayı bu farklılığı ortaya koyamama, yani, gerçekten red edilmesi gerekli olan bir hipotezin kabul edilerek II. tip hata yapma olasılığının artması durumu ortaya çıkar. Dolayısıyla varyansların homojen olmaması hata unsurlarında değişiklik meydana getirecektir.

Sezgin (1972) de bildirdiğine göre Cochran hataların homojen olmaması halinde denemeyi hataları homojen olan parçalara ayırıp her grubun kontrolü kendi hatasıyla yapılabileceğini söylemiştir.

Yapılan hatadaki (I. tip veya II. tip) değişmelerin ne kadar olacağı hakkında önceden bir fikir sahibi olmak mümkün değildir, fakat gruplardaki gözlem sayıları birbirine eşit ve her bir gruptaki tekerrür sayısının çok olduğu durumlardaki hata yapma ihtimali, örnek genişliklerinin eşit olmadığı ve az sayıda tekerrür içeren örneklerdeki hata yapma ihtimalinden daha düşük olacağı önceden bilinir (Düzgüneş vd 1987).

2.1.5. Varyans analizi tekniğinin ön şartları gerçekleşmediği takdirde karşılaşılan durumlar

Varyans analizi tekniğinin ön şartları gerçekleşmediği takdirde, analizi geçerli kılmanın çareleri olarak en iyi yöntemin ön şartları bozan durumlardan kaçınmak olduğu söylenebilir. Fakat bazen deneme materyalinin, bazende etkisi denenen muamelelerin yapısı yönünden bu imkansız olabilir. Kimya reaksiyonlarının hızı, organizmalarda büyüme olayı, bir kültürdeki bakteri yoğunluğunun değişmesi gibi olayların dayandığı ihtimal fonksiyonları normal olmayabilir (Sezgin 1972).

Özümleme miktarı, ışık yoğunluğu ve yaprak yüzeyinin toplanabilir değilde çarpımlı bir fonksiyonu olabilir (Sezgin 1972).

Bazı dağılımlarda, mesela poisson ve binom dağılımında ortalama ile varyans arasında bir ilişki vardır, bunlar birbirisinin fonksiyonudur. Transformasyon sonucunda transforme edilmiş değişkenler arasında da bu ilişki sonradan oluşabilir, söz konusu ilişki varyansları ortalamaya bağlı olarak değiştirdiği için varyansların homojenliği ön şartını bozar bu ön şartın tekrar yerine gelebilmesi için uygun transformasyon tekniği belirlenmelidir.

Bütün denemelerde seçilecek yol olayın matematik tanımına ve tarifine bağlıdır. Bazen, üzerinde denemeyi kuracak büyüklükte homojen materyal bulmak imkansız olabilir veya denemenin yürütülmesi esnasında elde olmayan sebeplerle bazı ünitelerde ön şartların tam olarak gerçekleşmesini engelleyen hasar ve etkiler olabilir. Bu nedenle analize başlamadan önce deneme hakkında geniş bir fikir edinmelidir.

Varyans analizi tekniğinin ön şartları gerçekleşmeden bu analize baş vurulursa hipotez kontrolünden önce kararlaştırılmış bulunan I. tip hata ihtimalinin hangi büyüklükte gerçekleştiğini bilmek mümkün olmaz. Yani araştırmacı H_0 hipotezini red ederken yanlış olma ihtimalinin en fazla % 5 olduğu kanısındayken

yanılma ihtimali belkide % 15, % 20 veya % 30 olarak gerçekleşmektedir. Böylesine bir durumun "karar verme" aşamasında ve buna bağlı olarak önerilecek uygulamalarda ne kadar büyük yanılırlara sebep olabileceği açıktır. Bu gibi durumlarda çözüm, bu çeşit ön şartları gerektirmeyen ve verilerin orjinal durumuna uygun başka bir analiz tekniği kullanmak mesela, parametrik olmayan testler gibi veya yeni bir analiz tekniği geliştirmektir. Parametrik olmayan testlerde, mukayese edilen gurupların temsil ettikleri populasyonların parametreleri ve dağılım fonksiyonu belli değildir veya belirtilmez. Bu testlerin uygulanmaları için normal bir dağılım şekli veya populasyon şartı aranmaz. Bu nedenlerdendir ki az gözlemlenilen deneylere uygulanabilirler ve nisbi etkinlikleri genellikle düşüktür.

Yukardaki bölümde açıklandığı üzere varyans analizi tekniğinin kullanılabilmesi için dört önemli ön şartın yerine gelmesi gereklidir. Bu ön şartların her birini tekrar ele alarak söz konusu ön şartın yerine gelmediği durumda kullanılan test istatistikindeki değişimler, yapılması gerekli düzeltmeler, daha uygun olan analiz teknikleri veya varyans analizi tekniğini kullanabilmek için söz konusu ön şartın yerine gelmesini sağlayacak uygun transformasyon tekniklerinden aşağıda özet halinde bahsedilmiştir.

2.2. Transformasyonlar

Varyans analizi tekniğini uygulamadan önce gözlemlerin dağılımının şekli normal değilse ve faktör etkileri eklenebilir değil ise bağımlı değişkenin doğrusal olmayan bir transformasyona tabi tutulması gerekir. Bu iki ön şart birbirini etkilemektedir. Yani eğer gözlemler normal dağılım göstermiyorsa eklenebilirlik ön şartıda sağlanmıyor demektir ve bu söylenilen durum tersi içinde geçerlidir.

Varyans analizinin ön şartlarından olan eklenebilirlik ön şartının sağlanabilmesi için yapılacak transformasyonun seçimi şu temellere dayanır.

- a) Toplanamayan etkiler için F değerin minimum olması, (Tukey 1949)
- b) İnteraksiyon etkisi için hesaplanan F değerin minimum olması,
- c) Hataya karşı test edilen muamele faktörünün etkisini gösteren F değerin maximizasyonu (Tukey 1950).

Uygulanacak transformasyondaki temel amaç ön şartları gerekli kılacak ve verilerin analizinden sonra yorumlanabilirliği sağlayacak bir transformasyon seçimidir.

Verilerin dağılımının şeklini değiştirmekle, asimetriklik bozularak kuyruk bölgesinde kayan gözlem sayısında azalma ve hataların küçülmesi de sağlanmaktadır. Bu değişimi sağlamak amacıyla bütün orjinal verilere bir matematiksel fonksiyon uygulayarak verileri transformasyona tabi tutmuş oluruz. Dağılımın şeklini değiştirmek için ortalamanın ve ölçeğin değişimi için uygulanan lineer transformasyonlardan daha fazlası gereklidir. Ortalama veya ölçek değişimi için uygulanan verilerin dağılımının şeklini etkileyen doğrusal transformasyonlardır. Dağılımın şeklini etkileyen güçlü transformasyonlara örnek olarak logaritmik ve karekök transformasyonları gösterilebilir.

Genellikle belirli bir (ön şartın) gerçekleşmesini sağlayan transformasyon aynı zamanda eş anlamlı olarak diğer ön şartların gerçekleşmesini de sağlar.

Verileri transformasyona tabi tutmanın nedenlerini özetle şu alt başlıklar halinde sıralayabiliriz.

- 1) Açıklama ve yorumu kolaylaştırmak,
- 2) Bir veri grubunda simetrikliği sağlamak,
- 3) Çeşitli veri gruplarındaki varyansların homojenliğini sağlamak,
- 4) İki değişken arasında doğrusal bir ilişki geliştirmek,
- 5) Basit bir toplama modeliyle verinin yapısını anlayabilecek duruma getirmektir.

Bir veri grubunun şeklinin simetrik olması istenen bir özelliktir. Çünkü veriler simetrik bir dağılımdan alınmış ise verilerin ortalamasının tahmini hesaplamalarda çok rahat kullanılır ve yapılan hesaplamaların hatası düşük, sonuçtaki yorumların anlaşılması kolay olur. Bağımlı değişkenlerin transformasyonlarında üzerinde durulan en önemli konulardan birisi varyans analizi için uygun koşullar sağlayacak şekilde varyansın sabitleştirilmesidir.

Varyansın sabitleştirilmesi amacıyla yapılan transformasyonlar genellikle normallik için en iyi transformasyonlar değildir (Snedecor, G.W. and Cochran W.G 1967).

Ancak burada hemen şu noktaya değinmekte fayda vardır. Uygulanan transformasyon eğer doğru bir şekilde tesbit edilmiş ise, bir çok yönden avantajlar sağlar. Fakat burada hemen şunu da eklemek gerekir ki doğru bir transformasyon metodu bile bir dezavantaj getirebilir. Uygulanan transformasyon bizi çoğunlukla az tanıdık bir ölçeğe taşır bu nedenle anlama ve yorumlama kabiliyetimizi kaybederiz. Mesela logaritmik transformasyon bu çeşit transformasyonlardan birisidir. Bu dezavantajlar, analizi yapana nazaran analizi değerlendirenlerle ilgilidir. Bu nedenle transformasyon uygulamak durumunda kaldığımız zaman kazançların gerekli olan çabayı karşılayıp karşılamadığını ve araştırmacı için ortaya çıkabilecek güçlükleri bilmemiz gereklidir.

Transformasyonun bir popülasyonun şeklini çok fazla değiştirip değiştiremeyeceğini, en büyük veri değerinin en küçük veri değerine oranlaması ile bulunabilir. Bu oran büyük olduğunda uygulanacak olan transformasyon yararlı olacaktır (yani > 20 olması istenir). Ama bu oran 2'den küçükse transformasyonun bir yararı olmaz (Hoyle 1973).

2.2.1. Güç transformasyonları

$X_1, X_2, \dots, X_n$ gibi bir veri grubunun transformasyonu her bir X_i değerini $T(X_i)$ ile yer değiştiren T fonksiyonu elde etmektir. Böylece söz konusu grubun dönüştürülmüş değerleri $T(X_1), \dots, T(X_n)$ olan bir fonksiyon olur.

Transformasyon tanımı

$$y = x^2 \text{ ve } y = \log x \dots\dots\dots (2.6)$$

gibi bilinen ve temel fonksiyonların önemli özelliklerine ihtiyaç duyan karmaşık fonksiyonlar içerir.

Miktar veya adet olarak elde edilen değerler söz konusu olduğunda uygulanacak olan transformasyon teknikleri aşağıdaki özelliklere sahiptir:

1) Bu transformasyon teknikleri populasyon içinde verinin sırasını korurlar yani artan fonksiyonlardır. Orjinal ölçekte büyük olan veri değerleri transformasyondan sonra yine büyüktür, ancak verilerin aralıkları değişebilir.

2) Veri noktaları arasında yeni değerler sokulmasından kaynaklanan küçük farklar vardır. Verilerin büyüklük sırası değişmediği için ortanca değerler yine ortanca değerlere dönüştürülmüş olur.

3) Bunlar sürekli fonksiyonlardır. Burada orjinal veri grubunda birbirine yakın değerler alan noktaların transformasyondan sonrada yine birbirine yakın olacağını ya da en azından kullanılmakta olan ölçekle alakalı olacağını garanti eder.

4) Bunlar düzgün fonksiyonlardır. Kullanılan fonksiyonlarının her düzeyinin türevleri vardır. Bu ihtiyaç fonksiyonların keskin köşelerinin olmamasını garanti eder.

5) Bunlar temel fonksiyonlarla kesin olarak belirtilmiştir. Böylece transformasyonlar en basit cep makinalarının yardımıyla çözülebilmektedir.

Bu özelliklerden sadece ilk bir kaçı toplama ve çarpma sabitlerinin seçimine bağlıdır.

a, b, c, d ve p gerçekte sayılar olmak üzere güç dönüşümleri aşağıdaki forma sahiptir:

$$T_p(x) = \begin{cases} ax^p + b & (p \neq 0) \dots\dots\dots (2.7) \\ c \log x + d & (p = 0) \end{cases}$$

1'den 5'e kadar olan şartların sağlanabilmesi için $p > 0$ için $a > 0$ ve $p < 0$ için $d < 0$ olması gerekir.

a, b, c ve d'nin değerlerinin seçimi tercihe bırakılmış ise genellikle uygunluk göz önüne alınarak ona göre bir seçim yapılır. Bunun içinde p'nin önemli olduğu ve verilerin analizinde bize yardımcı olacağı durumlarda p'ye göre bir seçim yapılır. 3 genel durum söz konusudur. Bunlar;

$$1) T_p(x) = \begin{cases} x^p & (p > 0) \\ \log x & (p = 0) \dots\dots\dots (2.8) \\ -x^p & (p < 0) \end{cases}$$

2) Transformasyonlar (dönüşümleri) biribiri ile karşılaştırmak ve onların matematiksel ve geometrik özelliklerini değerlendirmek istediğimizde şu şekilde özetleme yaparız.

$$T_p^*(x) = \begin{cases} (x^p - 1)/p & (p \neq 0) \dots\dots\dots (2.9) \\ \ln p & (p = 0) \end{cases}$$

ayrıca bu transformasyonları yukarıda tanımlanan beş özeliğide sunmaktadır (Manley 1976).

3) Transformasyona tabi tutulmuş veriler orjinal veriyi konum ve dağılım olarak andırması için bir veri gurubunu yeniden ifade etmek amacıyla sabitleri kıyaslama denilen bir sürece geçeriz. P'nin sabit tutulduğu ortamda a ve b sabitlerinin (veya P=0 iken c ve d sabitlerinin) seçimi diğer sabitlerin seçimine göre doğrusal transformasyonu temsil eder.

Örneğin;

$ax^P + b$ ifadesini, $Ax^P + B$ 'ye dönüştürmek için A/a çarpan ve $B - (A/a)b$ toplama sabiti alınarak dönüştürülür. Yani

$$\frac{A}{a} (ax^P + b) + (B - \frac{A}{a}b) = Ax^P + B \text{ elde edilir.} \dots\dots\dots (2.10)$$

Çünkü sabit olmayan doğrusal transformasyonlar yalnızca ortalamayı değiştirir. Ölçekte yeknesak bir değişim sağlar.

Güç transformasyonları ailesi pek çok bakımdan kullanışlıdır. Sabitler, yukarıda bahsedilen 3 yoldan belirlendiği zaman güç transformasyonları daha önce belirtilen 5 özelliği sağlar.

Varyansı sabitleştirmek için kullanılan genel metod varyans ve ortalama arasındaki ilişkiyi deneysel veya kuramsal yolla tesbit etmektir (Bartlett 1947).

Deneysel olarak grup ortalamasının logaritması alınarak bulunan bu değer karşısına grup varyansı bulunarak yazılır. Bir başka metod Bartlett'in kriterin de olduğu gibi varyansın heterojenliğini bir ölçüde minimize edebilmek için sınırlanmış grup içerisinde bir transformasyon seçmekten ibarettir.

2.2.2. Üstel transformasyon

Transformasyonun tanımına göre;

$$y = \begin{cases} (x^\lambda - 1)/\lambda & \lambda \neq 0 \dots\dots\dots (2.11) \\ \log_e(x) & \lambda = 0 \end{cases}$$

Burada eğer x negatif değerler alıyorsa λ parametresi iki tane olur ve transformasyon şu şekilde uygulanır.

$$y = \begin{cases} [(x^{\lambda_1} - 1)/\lambda_1] & \lambda_1 \neq 0 \dots\dots\dots (2.12) \\ \log_e(x + \lambda_2) & \lambda_1 = 0 \end{cases}$$

Üstel transformasyonlar ailesindeki parametre;

$$y = \begin{cases} \exp[(\gamma x) - 1]/\gamma & \gamma \neq 0 \dots\dots\dots (2.13) \\ x & \gamma = 0 \text{ tir.} \end{cases}$$

Negatif x değerleride dahil edildiğinde yukarıdakilere alternatif bir transformasyon olarak kullanılabilir (Manley 1976). Bu tip transformasyonlar eğik tek yönlü (Skew unmodel distributions) dağılımlar simetrik normal dağılıma dönüştürmek amacıyla kullanılır.

γ 'nın seçimi; γ 'yı seçtiğimizi ve y 'nin dağılımında normal olduğunu varsayalım x , $-\infty < x < +\infty$ arasında bir değer alıyorsa γ 'da pozitif bir değer almış ise y 'nin değişim sınırları $-1/\gamma$ ile $+\infty$ arasındadır. Eğer γ negatif olarak seçilmiş ise o zaman y 'nin değişim sınırları da $-\infty$ ile $-1/\gamma$ arasında olur. $\gamma = 0$ ise Y tam olarak normal dağılım gösterir. Bu durumda $Y = X$ olur.

Aynı durum (2.11) nolu transformasyonda da karşımıza çıkar. λ tek sayı ise veya sıfır ise bu durum transformasyon uygulayıcıları tarafından dikkate alınmalıdır.

Draper and Cox (1969) göstermiştir ki "eğer güç transformasyon parametresi λ olarak seçilmiş ise transformasyona uğrayan veri normal dağılım gösterir". Bu durum üssel dönüşümde γ 'nın iyi seçilmesi sonucunda maximum Likelihood tahmininin yapılabileceğini gösterir.

Veri değerlerini $x_1, \dots, x_n$ ile ve (2.12) nolu formülle transformasyona uğramış verileri $y_1, y_2, y_3, \dots, y_{n-1}, y_n$ ile gösterelim. Transformasyona uğramış y_i değerlerinin beklenen değerleri μ_1, μ_2, μ_n ve sabit varyansı ise σ^2 olsun, örneğe ait log likelihood şöyle tanımlanır.

$$L = \gamma \sum x_i - 1/2 \sum (y_i - \mu_i)^2 / \sigma^2 - n \log e[(2\pi)^{1/2} \sigma] \dots \dots \dots (2.14)$$

Eğer μ 'ler standart tipteki lineer model yardımıyla belirleniyorsa o zaman bu modelde yer alan parametrelerin Maximum Likelihood tahminleri alışlagelen regresyon tahminleri olur. σ^2 şöyle tahmin edilir.

$$\sigma^2 = \sum (y_i - \mu_i)^2 / n$$

γ^n nin maximum Likelihood tahmini ise şöyle bulunur.

$$\frac{\rho L}{\rho \gamma} = \sum x_i - \frac{1}{\gamma \sigma^2} \sum (y_i - \mu_i) [x_i(\gamma y_i + 1) - y_i] \dots \dots \dots (2.15)$$

Bu eşitlik sıfıra eşitlenir. Bu parametre (γ) için daha uygun bir gösterim sözü konusu değildir. μ_i 'lerin hepsi aynı ve μ 'ye eşit ise dönüşümden sonra örnek ortalamasının tahmini ve γ 'nin maximum likelihood tahmin edicisi şöyle hesaplanır.

$$\gamma^2 \approx 6 m^3 / (3m_2^2 - 7m_4) \dots \dots \dots (2.16)$$

m_r transformasyondan önce örnek ortalamasının r . momentidir. Bu yaklaşım $y_i = y(x_i)$ Taylor serisindeki X 'nin açılmasıyla elde edilir. Transformasyondan önceki ortalamayı bulmak için şu formül kullanılır.

$$y(x_i) = y(x + \delta x_i) = y(x) + \exp(\gamma x)(\delta x_i + 1/2 \gamma x_i^2 + 1/6 \gamma \delta x_i^3 + \dots) \dots \dots (2.17)$$

γ 'nın 3. veya daha yüksek dereceli kuvvetlerini dikkate almadan (2.13) nolu eşitliğin sağ tarafına (2.17) nolu eşitliği yerleştirerek hesaplama yapabiliriz. $\delta L/\delta = 0$ olma (2.16) nolu eşitlik elde edilir.

Transformasyondan sonraki ortalama ve varyans ise sırasıyla şöyle hesaplanır.

$$\mu = [y(x_i)/n = y(x) + \exp(\gamma x)(1/2\gamma m_2 + 1/6\gamma^2 m_3 + 1/24\gamma^4 m_4 + \dots)]$$

$$\sigma^2[y(x_i) - \mu]/n = \exp(2\gamma x)[m_2 + \gamma m_3 + \gamma^2(7/12m_4 - 1/4m_4) + \dots] \dots \dots (2.18)$$

(2.17) nolu eşitlikteki $\gamma \delta x_i$ 'nin pozitif değerlerinin kullanılmasıyla (2.16) nolu formüldeki yaklaşık değer elde edilir. (2.17) nolu seri transforme edilmiş y değişkeninin ortalamasına ilişkin üçüncü momentle ilişkilidir ve bu $m'_3 = \exp(3\gamma x)[m_3 - 1.5\gamma(m_2^2 - m_4) + \dots]$ ki bu durum eşitliğin yaklaşık sıfır olduğunu gösterir.

$$\gamma = 2 m_3/[3(m_2^2 - m_4)] \dots \dots \dots (2.19)$$

olur.

Bir çok tek yönlü (unimodal) dağılımlarda $m_4 \approx 3m_2$ olduğu durumlarda γ (2.16) ve (2.17) nolu denklemdeki γ 'lar aynı olur.

Pearson IV ve VI tipindeki dağılımların normal dağılıma dönüştürülmesinde üssel transformasyon kullanılır. "x" değerinin dağılımının diklik katsayısı en az 3 olan dağılım göstermesi durumunda uygulanacak üssel transformasyon diklik katsayısının 3'ten az olduğu durumdan daha başarılı sonuçlar vermektedir (Manley, 1976).

2.2.3. Ters transformasyonlar

Transforme edilen Y , orjinal değişkenlere X dersek ters transformasyon $Y = X^{-1} \dots \dots \dots (2.20)$ şeklinde tanımlanır.

Bu transformasyon çeşiti çok yaygın kullanılmaktadır ve ekstrem eğrilikleri düzeltmek amacıyla kullanılır. Elde edilen Y değerleri normallik, varyansların homojenliği, etkilerin eklenebilirliği geçerli olmasını sağlar.

2.2.4. Karekök transformasyon

$$y = (x + w)^{1/2}$$

Bu transformasyona, x ortalaması Θ olan poisson dağılımına sahip bir rastgele değişken olduğu zaman başvurulur. Ortalama ile varyans arasındaki ilişki bu transformasyon ile bozulur. 1930'larda ilk defa kullanılmıştır. Tippett (1934) $w = 0$ ile Bartlett (1936) problemi daha detaylı ele almış ve $w = 1/2$ olarak incelemiştir. Anscombe (1948), Johnson'un sonuçlarını kullanarak $w = 3/8$ 'in varyansı dengede tutmak için optimum olduğunu göstermiştir. Bu durumda;

$$\text{Var}(y) \approx 1/4(1 + 1/16\Theta^2)$$

$$\gamma_1(y) = 1/\Theta(1 + 7/16\Theta)$$

$$\gamma_2(y) = 1/\Theta(1 + 369/256\Theta)'dır. \dots\dots\dots (2.22)$$

Buradaki γ_1 ve γ_2 y'nin üçüncü ve 4. momentleridir ve sırasıyla eğrilik ve basıklığın ölçüleridir. Poisson dağılımında $\gamma = \sqrt{\Theta}$ ve $\gamma = \Theta^{-1}$ şeklinde hesaplanır. Daha sonraki çalışmalarda, Kihlberg, Herson, Schotz (1967). Kapsamlı bilgisayar çalışmasından $w = 0.386$ 'nın çoğu Θ değerleri için optimum bir değer olduğu sonucuna varmışlardır. Ayrıca $\Theta < 2$ olduğu durumlar da ise $w = 1/4$ 'ün en iyi sonuç verdiğini göstermişlerdir.

2.2.5. Logaritmik transformasyon

$$y = \log_e(x+w) \dots\dots\dots (2.23)$$

Bu transformasyon 3 şekilde ortaya çıkar. İlk olarak lognormal dağılım olarak adlandırılan dağılımda, örneğin bir ülkenin populasyonundaki gelir dağılımı bu tip bir dağılımdır, yukarıdaki tanımlanan transformasyon kullanılır. Eğer y normal dağılıyor ise x 'in lognormal dağıldığı söylenebilir.

Bartlett, Kendall (1946), Anscombe (1948), Klezowski (1949), Moore (1958) ve Healy (1967) geriye kalan üç varsayımın varyansların homojenliği, etkilerin eklenebilirliği, gözlemlerin bağımsızlığı sağlanması için bu transformasyonu deneysel kullanılması gerektiğini açıklamışlardır. Bu transformasyon varyansı stabilize etmek amacıyla da kullanılır eğer $\text{Var}(x) = D^2(\Theta) = \Theta^2$ ise veya x'in standart sapması Θ 'ya oranlanırsa o zaman $g(x) = y = \log(x)$ varyansı dengede tutan uygun transformasyon olur.

2.2.6. Arcsin transformasyon

$$y = \text{Arcsin}(x) \dots\dots\dots (2.24)$$

Bu transformasyon ilk olarak Fisher (1922) tarafından önerilmiştir. Yine Fisher (1930) binomiyal değişkenlerin varyanslarını sabitleştirmek için kullanılmıştır.

2.2.7. Hiperbolik transformasyon

Bu transformasyon x rastgele değişkeni negatif binomiyal dağıldığı zaman varyansı sabitleştirmek için kullanılır.

$$\text{Prob}(x=z) = \frac{r(x+k)}{x!r(k)} \left(\frac{\Theta}{\Theta+k}\right)^x, \quad x = 0,1,2,\dots \dots\dots (2.25)$$

Beall (1942) bu transformasyonu $w_1 = w_2 = 0$ olarak ilk defa ortaya atan kişidir. Anscombe (1948) bu transformasyonu $w_1 = 3/8$ ve $w_2 = 3/4$ alınmasını teklif etmiştir.

Bu durumda büyük Θ değerleri için ve sabit k/Θ için $\text{Var}(Y) = 1/4 + \Theta(1/\Theta^2)$ ve $Y = \log(x + 1/2k)$ nın daha basit fakat daha güçlü sonuç verdiğini açıklamıştır.

2.2.8. Açısal transformasyon

$$y = \arcsin \sqrt{\frac{x+w_1}{n+w_2}} \dots\dots\dots (2.26)$$

Bu transformasyonun kullanım amacı, binomiyal bir değişkenin varyansını stabilize etmektir. Bartlett (1936), $w_1 = 1/2$ ve $w_2 = 0$ olduğu durumda stabil bir $\text{var}(Y)$ ortaya çıktığını belirtmiştir.

$$y = \arcsin \sqrt{\frac{(x+3/8)}{(n+3/4)}} \dots\dots\dots (2.27)$$

Bu durumda;

$$\text{Var}(Y) = 1/4 + \Theta(1/n^2) \dots\dots\dots (2.28)$$

$$\gamma_1(Y) \approx \frac{2\Theta-n}{2(n\Theta(n-\Theta))^{1/2}}$$

$$\gamma_2(Y) \approx \frac{n^2-2\Theta(n-\Theta)}{n\Theta(n-\Theta)} \dots\dots\dots (2.29)$$

$x = 0$ için veya $x = n$ olduğu zaman

$$y = \arcsin \sqrt{1/4n} \quad x = 0$$

$$y = 90^\circ - \arcsin \sqrt{1/4n} \quad x = n \text{ kullanılır.} \dots\dots\dots (2.30)$$

2.2.9. İhtimal integral transformasyonu

$$y = \int_{-\infty}^t P(x) dx \dots\dots\dots (2.31)$$

Bu transformasyon ilk defa Fisher (1932) ve Karl-Pearson tarafından kullanılmıştır. Buradaki $P(x)$, x 'in ihtimal yoğunluk fonksiyonudur. y değişkeni bu durumda $(0,1)$ aralığında uniform dağılır. Bu yüzden bu transformasyon etkilemez. $\text{Var}(Y) = 1/12$ 'de mükemmel bir şekilde varyansı stabilize eder. David and Johnson (1948, 1950) ve David (1950) $P(x)$ 'in bazı bilinmeyen parametrelere bağlı olduğu durumlarda bu transformasyonu daha detaylı bir şekilde incelemişlerdir.

Eğer grup ortalamaları birbirine eşit buna karşılık grup varyansları birbirinden farklı ise o zaman bu durum H_0 hipotezinin red edilme ihtimalini artırır. Büyük serbestlik derecesine sahip olan toplanmış varyans grup varyanslarının tahminlerine ait ortalamanın en iyi tahminidir. Bağımlı değişken y olsun ve hata varyansı ise ortalamanın bir fonksiyonu olan $f(y)$ olsun. Örneğin, eğer y binomiyal bir değişken ise ve 10 birimlik bir örneğin ortalaması $10p$ olur. Buradaki p istenen olayın örnekteki oluş ihtimalidir. Varyans ise $10p(1-p)$ olur. Burada $f(p) = np(1-p)$ 'dir. $g(\mu) = k\mu(1-\mu)$ şeklindeki binomiyal dağılım gösteren veriler için uygulanacak transformasyon şöyle tanımlanır.

$$h(x) = \text{Sin}^{-1}(x)^{1/2} \text{ dir.} \dots\dots\dots (2.32)$$

Yani $\text{Arcsin}/\sqrt{p}$ transformasyonu uygun görülür. h 'in ölçümü radyan ise büyük örneklerde h 'in varyansı yaklaşık olarak $(1/4n)$ 'e eşit olur. Bu transformasyon orta derecede değişen oranlar için en uygun transformasyondur.

Anscombe (1948) (2.32) nolu transformasyon denklemine basit değişiklik yaparak yeniden şöyle tanımlamıştır.

$$h^*(x) = \text{Sin}^{-1}(p^*)^{1/2} \text{ buradaki} \dots\dots\dots (2.33)$$

$$p^* = \frac{np + 3/8}{n + 3/4}, \text{ dir.}$$

Bu düzeltmenin amacı transformasyondan önce oranların 0.50'ye yaklaşmasını sağlamaktadır. Eğer np ve nq'un en küçük değeri 10 oluyorsa $h(x) = \text{Sin}^{-1}(x)^{1/2}$ transformasyonu $h^*(x) = \text{Sin}^{-1}(p)^{1/2}$ transformasyonu ise $\min(np, nq) > 2$ olduğu durumda daha iyi sonuç verir.

Bu transformasyonun etkisi en düşükten en yükseğe kadar olan yüzdeler içinde geçerlidir. % 1 ile % 2 arasındaki farkların ölçüsü % 50 ile % 54 arasındaki farklara eşittir. $g(\mu) = k\mu^2(1-\mu)^2$ fonksiyonuna ait deneysel verilere uygulanacak transformasyon ;

$$h(x) = \ln \left[\frac{x}{1-x} \right] \text{şeklindedir.} \dots\dots\dots (2.34)$$

Bu transformasyon logaritmik tam sayı ile artan transformasyon olarak bilinir ve sınıflandırılmış veriler için kullanılır.

$g(\mu) = k(1-\mu^2)^2$ fonksiyonuna ait örnek korrelasyon katsayısına uygulanacak transformasyon şöyle tanımlanır.

$$h(x) = \ln \left[\frac{(1+r)}{(1-r)} \right] \dots\dots\dots (2.36)$$

Bu transformasyon klasik fisher Z transformasyonu olarak bilinir ve iki değişkenli normal dağılımdan alınan örnekten hesaplanan korrelasyon katsayısı için uygulanır.

$g(\mu) = k\mu$ (poisson dağılımı) gösteren deneysel verilere uygulanacak transformasyon

$$h(x) = (x)^3 \text{dir.} \dots\dots\dots (2.35)$$

Poisson dağılımının diğer istatistik dağılımlarla ilgili olan birkaç önemli ortak özelliği vardır. Poisson dağılımında meydana gelen sonuçlar zaman içerisinde rastgele oluşur ve bu sonuçların sayısı sabit aralıklarla gerçekleşir. Eğer np büyük ise (20 ve daha fazla ise) bu durumda poisson dağılımı normal dağılıma yaklaşır ve transformasyona ihtiyaç kalmaz. Eğer np küçük ise varyansın stabilizasyonu için

küçük bir değişiklik yapılmalıdır (Anscombe 1948); (Freeman and Tukey 1950) bu konu üzerinde tartışmışlar ve gerekli olan transformasyonları sırasıyla aşağıdaki şekilde tanımlanmışlardır.

$$h^*(x) = (x + 3.8)^{1/2} \dots\dots\dots (2.36)$$

$$h^{**}(x) = (x)^{1/2} + (x+1)^{1/2} \dots\dots\dots (2.37)$$

$\mu < 2$ için (2.37)'da tanımlanan transformasyon daha uygundur. Buna karşılık diğer değerler için bu iki transformasyon arasında çok küçük farklar vardır (Smith, J.E Keith, 1976).

Çizelge 2.1'de bu transformasyonların özeti verilmiştir.

Çizelge 2.1. Transformasyonların özeti

μ cinsinden varyans	Transformasyon	Yeni verilerdeki yaklaşık varyans	Dağılımın şekli
a) μ b) $\lambda^2\mu$	$\sqrt{y}$ veya küçük gözlemler için $\sqrt{y+0.5}$	0.25 $0.25 \lambda^2$	Poisson Amprik
$2\mu^2/(n-1)$	$\ln y$	$2/(n-1)$	Örnek varyansları
$\lambda^2\mu^2$	$\log_{10}y, \log_{10}(y+1)$ $\ln y, \ln(y+1)$	$0.189\lambda^2$ λ^2	Amprik
$(\mu(1-\mu)/n)$	$\sin^{-1}\sqrt{y}$ (Radyan) $\sin^{-1}\sqrt{y}$ (Derece)	$0.25 n$ $821/n$	Binomiyal
$(\mu(1-\mu)/n)$	Probit	Sabit değil	
$(\mu(1-\mu)/n)$	$\frac{y}{\ln(\frac{y}{1-y})}$	$1/(\mu n(1-\mu))$	
$\lambda^2\mu^2(1-\mu)^2$	$\frac{y}{\ln(\frac{y}{1-y})}$	λ^2	Amprik
$((1-\mu^2)^2/(n-1))$	$\frac{1}{2} \ln(\frac{1+y}{1-y})$	$\frac{1}{n-3}$	Korelasyon Katsayısı
$\mu + \lambda^2\mu^2$ $\gamma^2(\mu + \lambda^2\mu^2)$	$\sinh^{-1}(\lambda\sqrt{y})$ $\sinh^{-1}(\lambda\sqrt{y+0.5})$ (Küçük sayılar için)	0.25	Negatif Binomiyal

2.3. Kaynak Araştırması

2.3.1. Genel açıklamalar

Veri transformasyonu istatistik metodlarının önemli konularından birisidir. İstatistik tekniklerinin temelinde doğrusal ve doğrusal olmayan bir model söz konusudur. Bu modelde etkisi araştırılan ve tesadüfe atfedilen faktör etkileri yer alır. Birçok durumda, üzerinde çalışılan model birtakım varsayımların yerine geldiği takdirde kullanılabilir. Böyle bir durumda verilerin yapısı söz konusu olur. Verilerin yapısı uygulanacak istatistik tekniğe uygun düşmediği takdirde bu yapı üzerinde düşünmek gerekir. Bu tip problemler üzerinde uzun yıllar çalışılmış ve birçok konuya açıklık getirilmiş olup, çıkan sorunlar içinde çözüm aranmaktadır.

Tezde, bahsedilen istatistik tekniklerinden varyans analizi tekniği ele alınmış olup bu tekniğin uygulanabilmesi için gerekli önşartlar açıklanarak elde edilecek verilerin yapısında aranması gerekli özellikler üzerinde durulmuştur. Bu konuyla ilgili günümüze kadar yapılan çalışmalarla ilgili özet bilgiler aşağıda verilmiştir.

Hoyle (1973) verileri transformasyona tabi tutmanın sebepleri, uygun transformasyon tekniği geliştirmenin yolları ve yaygın kullanılan transformasyon tekniklerini ele almıştır. Ayrıca orjinal verilerle transforme edilmiş verileri karşılaştırmalı olarak ele almak amacıyla model varsayımlarından bir kısmı yerine gelmediği takdirde uygulanan istatistik tekniğinin sonuçlarındaki değişimleri incelemiştir.

Manly (1976), Box and Cox (1964)'un geliştirdiği güç transformasyonlarına alternatif olabilecek üstel transformasyon tekniğinden bahsetmiştir. Negatif veriler söz konusu olduğu zaman bu transformasyon tekniğinin daha kullanışlı bir teknik olduğunu açıklamış ve özellikle gamma dağılımı gibi uzun kuyruk bölgesine sahip dağılımları normal dağılıma dönüştürmek amacıyla üstel transformasyon tekniği kullanmanın daha uygun sonuçlar vereceğini tartışmıştır. Fakat, J ve U şeklindeki

dağılımları normal dağılımı dönüştürmek için çok etkili olmayan bir transformasyon tekniği olduğunu da vurgulamıştır.

Yapılan bu çalışmada regresyon veya varyans analizi tekniklerinde hata varyansının homojen olmadığı durumlarda kullanılacak en uygun transformasyon tekniğinin $\exp(-2\gamma x)$ olan üstel transformasyon olduğu açıklanmıştır.

Draper and Cox (1969) J ve U şeklindeki dağılımları normal dağılıma dönüştürmek amacıyla en etkili transformasyon tekniğinin Box-Cox (1964) tarafından geliştirilen güç transformasyonları olduğunu tartışmışlardır. Ayrıca Draper ve Hunter (1969), Linsey (1975) ve Hinkley (1975) güç transformasyonları ailesi üzerinde durmuş ve birçok durumlarda güç transformasyonlarının daha etkili olduğunu vurgulamışlardır.

Gaudry and Laferriere (1989), lineer regresyon modellerinde, bağımlı değişkenin Box-Cox metodu ile transformasyonu üzerinde durmuşlardır.

Ermerson and Stoto (1982) güç transformasyonları ailesinden en yaygın kullanılan X^2 ve logaritmik transformasyonlar üzerinde çalışmışlar ve özellikle sayılarak elde edilen verilere uygulanabilecek transformasyonları incelemişlerdir. Ayrıca verileri transformasyona tabi tutmanın sebepleri üzerinde durmuşlardır. Simetrik dağılım gösteren verilerde, ortalama ile varyans arasında ilişki olduğu zaman, bu ilişkiyi bozarak homojen bir varyans elde etmek için uygulanacak transformasyon çeşitlerini açıklamışlardır. Gruplardaki varyansların homojen olduğu koşullarda logaritmik veya güç transformasyonlarından uygun λ parametresi belirlenerek uygulanacak transformasyon tekniklerini incelemişlerdir.

Schiesselman (1973, 1974) iki yönlü varyans analizi modellerindeki varsayımların bozulduğu durumlarda uygulanabilecek güç transformasyonlarını ele almıştır. Örnek ortalaması ile varyansı arasındaki ilişkiyi tespit etmek için örnek ortalamasının logaritmasının örnek varyansının logaritmasına regresyon katsayısını t-testi ile kontrol edilmesi gerektiğini vurgulayarak t-testi sonucunda H_0 hipotezinin

red edilmesinin söz konusu ilişkinin varlığını göstereceğini ve bu durumda güç transformasyonlarına başvurabileceğini vurgulamıştır. Ayrıca normallik ve eklemelilik varsayımlarının bozulduğu durumlarda ele almıştır.

Atkinson (1973), Box ve cox transformasyon yöntemi ile tahmin edilen transformasyon parametresi hakkında yapılan testler ve varyans analizi tekniğinin önşartları ve özellikle normallik homojen bir varyans ile ilgili önşartların bozukluğu durumlarını incelemişler ve lineer ve lineer olmayan transformasyonlar hakkında açıklamalarda bulunmuştur.

Box and Cox (1964) varyans analizi modellerinde gerekli önşartların yerine gelmediği durumlarda orijinal gözlemlere uygulanacak transformasyon tekniğini araştırmışlardır. Ayrıca modelde yer alan parametrelerin likelihood fonksiyonu yardımıyla tahmini üzerinde de durmuşlardır.

Spitzer (1984) lineer modelde yer alan parametrelerin varyans-kovaryans matrisinin Box-Cox transformasyonu uygulandıktan sonra nasıl etkilendiğini araştırmıştır.

Rubin (1984) transformasyona tabi tutulan verilerin analizi üzerinde durmuş, ayrıca verilerin normal dağılması amacıyla uygulanan transformasyon sonucunda elde edilen tahmin değerleriyle transformasyon öncesi tahmin değerlerini karşılaştırmıştır.

Bickel and Doksum (1981) Box-Cox transformasyonunu ele alarak normallik varsayımının bozulduğu durumlarda bu transformasyon tekniğinin gücünü tartışmışlardır. Özellikle uzun kuyruk bölgesine sahip dağılımlar için uygun transformasyon metodunu araştırmışlardır. Emerson (1982) sayılarak elde edilen verilere uygulanacak güç transformasyonu üzerinde durmuş, varyans analizi tekniğinin önşartlarını dikkate alarak uygulanması gerekli transformasyon tekniklerinin matematiksel kökeni üzerinde araştırma yapmıştır.

Spitzer (1978) örnekleme denemelerinde küçük örnek genişliklerini ele alarak bağımlı ve bağımsız değişkenlere uygulanacak Box-Cox transformasyonunun etkisini araştırmıştır. Aynı transformasyon tekniğinin önemli bir problem oluşturmadığını ancak transformasyon parametresi olan λ 'nın işaret ve büyüklüğünün tahminlerin varyansını ve bağımlı değişkenin varyasyon katsayısını etkilediğini açıklamıştır.

Spitzer (1981) Box-Cox metodu kullanılan modellerdeki parametrelerin tahmininde likelihood fonksiyonu kullanmıştır.

Kortler (1950), Thoni (1963b) özellikle biyolojik veriler için logaritmik transformasyonun hataların homojen varyansa sahip olmasını sağladığını bildirmişlerdir.

Smith (1976) varyans analizi modellerinde bağımlı değişkene transformasyon uygulamanın sebeplerini araştırmıştır. Temel olarak ele aldığı konular, verilerin dağılımının şekli ve varyansların homojen olmama durumudur. Ayrıca veri yapısındaki farklılıklar üzerinde de durmuştur.

Tukey (1955) özellikle transformasyon ailesinden ziyade tek bir transformasyon üzerinde yoğunlaşmış ve transformasyonların yapısını karşılaştırmıştır.

Spitzer (1982) Box-Cox transformasyonu uygulanan modellerde yer alan parametrelerin tahmini için düzenlenmiş Newton algoritması üzerinde durmuştur. Tahminde kullanılan metod çok olabilirlik tahmin edicisidir. Düzenlenmiş algoritmanın etkinliğini diğer tahmin teknikleri ile karşılaştırmış ve bu algoritma ile elde edilen parametre tahminlerinin diğer tekniklerle elde edilen parametre tahminleriyle aynı sonucu vermesine karşın Newton algoritmasındaki tahmin değerleri iki veya dört kat daha hızlı elde edilmiştir.

Zeki (1979) verilerin normal dağılımdan sonra durumlarını ve test istatistiğine etkilerini incelemiştir.

Sezgin (1972) varyans analizi tekniğinin önşartları yerine gelmediği durumlarda uygulanması gerekli transformasyon tekniklerini incelemiştir. Önşartlardan normallik ve eklenebilirlik durumunu ele alarak bu önşartların bozulduğu durumlar, test sonuçlarına etkileri ve uygun transformasyon tekniklerini incelemiştir.

Ergen (1979) varyans analizi tekniğinin önşartları gerçekleşmediği durumlarda Box-Cox transformasyon tekniğinin uygulama şeklini açıklamıştır. Uniform dağılım gösteren verilerde transformasyon parametresinin tahmin değeri $\lambda = -1.07574$ olarak bulunmuş ve varyanslar heterojen olduğu durumda bu değer ters transformasyonun homojenliği sağlamak için çok etkili olacağını söylemiştir. Üslü verilerde varyans heterojen iken transformasyon parametresinin tahmin değerini $\lambda = -1.17981$ olarak bulmuş ve yine ters transformasyon ile varyansların homojenlik önşartının sağlanabileceğini vurgulamıştır. Ayrıca bu transformasyon çeşidinin sınırlı poisson dağılımı gösteren veriler için de geçerli olduğunu göstermiştir. Normal dağılım gösteren verilerde ise varyansların heterojenlik durumu için hesaplanan transformasyon parametresinin tahmin değeri $\lambda = -0.09338$ 'dir. Bu durum eğer varyanslarla ortalama arasında bir ilişki varsa logaritmik transformasyonla bu ilişkinin bozularak varyansların homojenlik önşartının sağlanabileceğini göstermiştir. Ancak bu son açıklanan kısım varyansın büyük bir değer aldığı durumlar için geçerlidir. Varyanslar küçük ise herhangi bir transformasyon uygulamanın heterojenliği etkilemediği açıklanmıştır. Ayrıca bu çalışmada normallik ve eklenebilirlik önşartları üzerine de ayrıntılı açıklamalarda bulunulmuştur.

Çetingöz (1980) soya fasulyesinde çeşitli özellikler dikkate alarak elde edilen verilerde varyans analizi tekniği uygulamadan önce normallik, eklenebilirlik ve varyansların homojenlik varsayımlarını incelemiş ve bu varsayımların bozulduğu durumlarda uygun transformasyon tekniğini Box-Cox metodu ile belirleyerek sonuçlar üzerine etkisini araştırmıştır.



3. MATERYAL VE METOD

3.1. Materyal

Bu çalışmada, materyal olarak bilgisayarda üretilmiş tesadüf sayıları kullanılmıştır. Bu sayılardan bir kısmı aynı ortalamalı ve değişik varyanslı normal dağılım gösteren populasyonlardan, bir diğer kısmı ise 1, 3, 5 ve 15 serbestlik dereceli χ dağılımı gösteren populasyonlardan alınmıştır. Bu dağılımlardan alınan gözlemler iki, üç ve dört grup halinde düzenlenmiş ve her grupta değişik örnek genişlikleri ele alınmıştır.

Tesadüfi sayı üretimi ve hesaplamalar için QBASIC 4.0 programlama dilinden yararlanılmıştır.

3.2. Metod

Varyansların heterojenliğinin varyans analizi tekniğinde hipotez kontrolü yaparken 1. tip hata yapma ihtimali olan α üzerine etkisini araştırmak amacıyla materyal bölümünde açıklanan verilerle yapılan simülasyon çalışmasında varyansların heterojenliğinin çeşitli seviyeleri ile gruplardaki gözlem sayılarının çeşitli seviyeleri ele alınarak grup sayısına ait tüm kombinasyonlarda varyans analizi tekniği uygulanmış, bu denemeler her kombinasyon için 100.000 defa tekrarlanmıştır. Her bir denemeden elde edilen sonuçlar ayrı ayrı bulunmuş ve bu deneme sonuçları toplu halde değerlendirilmiştir. Hesaplamalar sonucunda orijinal verilerde red edilen H_0 hipotezi sayıları ve aynı verilere logaritmik, karekök ve ters transformasyon uygulandığında red edilen hipotez sayıları tespit edilmiştir.

Yapılan simülasyon çalışmasında, ortalaması aynı, varyans oranları aşağıda belirtildiği gibi değişen iki farklı normal dağılım gösteren populasyondan gruplardaki gözlem sayısı eşit ve farklı olan denemelerde ayrı ayrı olmak üzere toplam 50 kombinasyona ait hesaplamalar yürütülmüştür. Bu kombinasyonlarda

varyans oranları 1:1 ile 1:10 arasında, örnek genişliği 2 ile 30 arasında değişmektedir.

Üç farklı populasyon için gruptaki gözlem sayıları eşit iken varyans oranları 1:1:1 ile 1:25:50 arasında, örnek genişlikleri ise 2 ile 30 arasında değişen toplam 20 kombinasyon, gruptaki gözlem sayıları farklı olduğunda varyans oranları 1:1:1 ile 1:4:8 arasında, örnek genişlikleri 2 ile 50 arasında değişen 16 kombinasyon üzerinde hesaplamalar yürütülmüştür.

Dört farklı populasyon için grupta gözlem sayıları eşit olduğunda varyans oranları 1:1:1:1 ile 1:25:50:75 arasında, gruptaki gözlem sayıları ise 2 ile 30 arasında değişen 20 kombinasyon, grupta gözlem sayıları eşit olmadığında varyans oranları 1:1:1:1 ile 1:5:7:9 arasında, örnek genişlikleri ise 20 ile 50 arasında değişen 16 kombinasyon dikkate alınarak hesaplamalar yürütülmüştür.

Benzer şekilde, gruptaki gözlem sayılarının eşit ve farklı olduğu durumlar için ve ayrıca gözlem sayısının az ve çok olduğu durumlar için χ^2 dağılımı gösteren verilere uygulanan logaritmik, karekök ve ters transformasyonların, varyans analizi tekniği sonucunda 1.tip hata yapma ihtimali üzerine etkisini araştırmak amacıyla χ^2 dağılımından alınan tesadüf örneklerinin herbirisi iki, üç ve dört grup oluşturacak şekilde düzenlenmiştir. Tezde, yukarıda bahsedilen transformasyonların etkisini araştırmak amacıyla bir, üç, beş ve onbeş serbestlik dereceli χ^2 dağılımları kullanılmıştır.

Bir serbestlik dereceli χ^2 dağılımından alınan örnekler önce iki gruba ayrılmış ve gruptaki gözlem sayıları 2:2, 5:5, 15:15 ve 30:30 olacak şekilde düzenlenmiştir. Böylece iki gruptaki gözlem sayılarının eşit olduğu koşullar için varyans analizi tekniği ile sonuçlar değerlendirilmiştir. Bir serbestlik dereceli χ^2 dağılımından alınan örnekler 3 grup halinde düzenlendikten sonra ilk olarak gruptaki gözlem sayılarının eşit olduğu farklı koşullar ele alınmıştır. Bu koşullardaki örnek genişlikleri ise sırasıyla şöyledir. 2:2:2, 5:5:5, 15:15:15 ve

30:30:30. Daha sonra 3 gruptaki gözlem sayıları birbirinden farklı tutulmuş ve bu amaçla örnek genişlikleri 2:5:10, 5:10:15, 5:10:20, 10:20:50 şeklinde düzenlenmiştir. Grup sayısı 4 olarak alındığı zaman guruplardaki gözlem sayılarının eşit olduğu durumlar için belirlenen örnek genişlikleri 2:2:2:2, 5:5:5:5, 15:15:15:15 ve 30:30:30:30 şeklindedir. Gruplardaki gözlem sayılarının farklı olduğu durumlar için belirlenen örnek genişlikleri ise 2:4:6:8, 5:10:15:20, 5:10:20:40 ve 15:20:30:50 şeklindedir.

χ^2 dağılımında, serbestlik derecesinde meydana gelen değişmeler neticesinde varyans analizi tekniği ile elde edilen sonuçların nasıl değiştiğini görmek amacıyla bir serbestlik dereceli χ^2 dağılımına ilaveten 3, 5 ve 15 serbestlik dereceli χ^2 dağılımlarından alınmış tesadüf örneklerinden de yararlanılmıştır, 3, 5 ve 15 serbestlik dereceli χ^2 dağılımları için yapılan düzenlemeler, belirlenen grup sayıları ve gruplardaki gözlem sayıları 1 serbetlik dereceli χ^2 dağılımı için anlatılan durumlar gibidir. Böylece χ^2 dağılımı gösteren verilere uygulanacak varyans analizi tekniğinde sonuçların ne şekilde değiştiğini, 1. tip hatadaki değişimin ne yönde olduğunu görmek amacıyla eşit örnek genişlikleri için toplam 48 farklı koşul, farklı örnek genişlikleri için ise toplam 56 farklı koşul değerlendirilmiştir. Bu koşulları değerlendirmek amacıyla orijinal verilere karekök, ters ve logaritmik transformasyon uygulanmış ve varyans analizi tekniği ile I. tip hata yapma ihtimalleri belirlenmiştir. Ayrıca denemede oluşturulan kombinasyonlar 2, 3 ve 4 grup halinde düzenlenmiştir.

Normal ve χ^2 dağılımlı populasyonlardaki grup sayısı, varyans oranları, eşit ve eşit olmayan örnek genişliklerine ilişkin bilgiler Çizelge 3.1 ve 3.2’de topluca verilmiştir. Bu çizelgelerdeki bütün kombinasyonları için simulasyon çalışmaları orijinal ve transforme edilmiş gözlemlerde 100.000’er defa tekrarlanmıştır.

Çizelge 3.1. Normal dağılımlı populasyonlara ilişkin simulasyon şeması

		Örnek genişlikleri	Varyans oranları
2 Grup	Eşit	2:2, 5:5, 15:15, 30:30	1:1, 1:1.5, 1:3, 1:5, 1:10
	Farklı	2:5, 2:10, 2:20, 5:10, 2:20, 10:30	1:1, 1:1.5, 1:3, 1:5, 1:10
3 Grup	Eşit	2:2:2, 5:5:5, 15:15:15, 30:30:30	1:1:1, 1:1.5:2, 1:3:5, 1:4:8, 1:25:50
	Farklı	2:5:10, 5:10:15, 5:10:20, 10:20:50	1:1:1, 1:1.5:2, 1:3:5, 1:4:8
4 Grup	Eşit	2:2:2:2, 5:5:5:5, 15:15:15:15, 30:30:30:30	1:1:1:1, 1:1.5:2:2.5, 1:3:5:7, 1:3:5:9, 1:25:50:75
	Farklı	2:4:6:8, 5:10:15:20, 5:10:20:40, 15:20:30:50	1:1:1:1, 1:1.5:2:2.5, 1:3:5:7, 1:5:7:9

Çizelge 3.2. χ^2 dağılımlı populasyonlara ilişkin simulasyon düzenleri

Örnek genişlikleri			
Grup Sayısı	Serbestlik derecesi	Eşit	Farklı
2	1, 3, 5, 15	2, 5, 15, 30	2:5, 2:10, 2:20:5:10, 5:20:10:30
3	1, 3, 5, 15	2, 5, 15, 30	2:5:10, 5:10:15, 5:10:20, 10:20:50
4	1, 3, 5, 15	2, 5, 15, 30	2:4:6:8, 5:10:15:20, 5:10:20:40, 15:20:30:50

4. SONUÇLAR VE TARTIŞMA

Bu çalışmada, materyal ve metod bölümünde açıklanan, normal dağılım gösteren veriler için varyansların heterojen olduğu durumlarda varyans analizi tekniği sonuçların ne şekilde etkilendiği açıklanmış ve varyansların homojenlik önşartı sağlanmadığı durumlarda verilere logaritmik, ters ve karekök transformasyonları uygulamanın etkili olup olmadığı araştırılmıştır. Ayrıca χ^2 dağılımı gösteren veriler için varyans analizi tekniğinin uygulamasına bakılmış yukarıda sayılan transformasyonların sonuçları ne şekilde etkilediğini göstermek amacıyla yapılan simülasyon çalışması sonuçları tablo halinde sunulmuştur.

4.1. Normal Dağılım Gösteren Veriler

4.1.1. İki grup ve gruplardaki gözlem sayıları eşit iken elde edilen sonuçlar

Bu başlık altında ele alınan tesadüf sayıları iki grup halinde düzenlendikten sonra bu grupların varyanslarının farklı durumları için varyans analizi tekniği uygulanmış ve gerçekleşen 1. tip hata olasılıkları saptanmıştır. Hesaplama sonuçları Çizelge 4.1'de toplu olarak verilmiştir.

Çizelge 4.1'den görüleceği üzere, gruplardaki varyanslar homojen olduklarında yapılan varyans analizi sonucunda gerçekleşen 1. tip hata olasılıkları, örnek genişliklerinin 2, 5, 15, 30 olduğu durumlarda, sırasıyla, % 5.1, % 4.9, % 5.0, % 5.0 ve % 4.9 olarak gerçekleşmekte olup, bunların başlangıçta kararlaştırılmış olan % 5.0'den pek farklı olmadıkları gözlenmektedir.

Varyans oranları 1:3 olduğunda ise gerçekleşen 1. tip hata olasılıklarında, özellikle örnek genişlikleri 2 ve 5 olduğunda, bir miktar artış gözlenmektedir. Örnek genişlikleri 15 ve 30 olduğunda ise gerçekleşen 1. tip hata olasılıkları kararlaştırılmış olan % 5.0 civarında kalmaktadır.

Gruplardaki varyans oranları 1:5 olduğunda ise, örnek genişlikleri 2, 5 ve 15 olduğunda gerçekleşen 1. tip hata olasılıkları, % 6.4, % 6.1 ve % 5.4 olmaktadır. Burada örnek genişliği 30 olduğunda gerçekleşen 1. tip hata olasılığı, başlangıçta kararlaştırılmış olan % 5.0'e yaklaşmaktadır.

Çizelge 4.1. Grup sayısı 2 ve gruplardaki gözlem sayıları eşit olduğunda 100000 deneme sonucunda red edilen hipotez sayıları

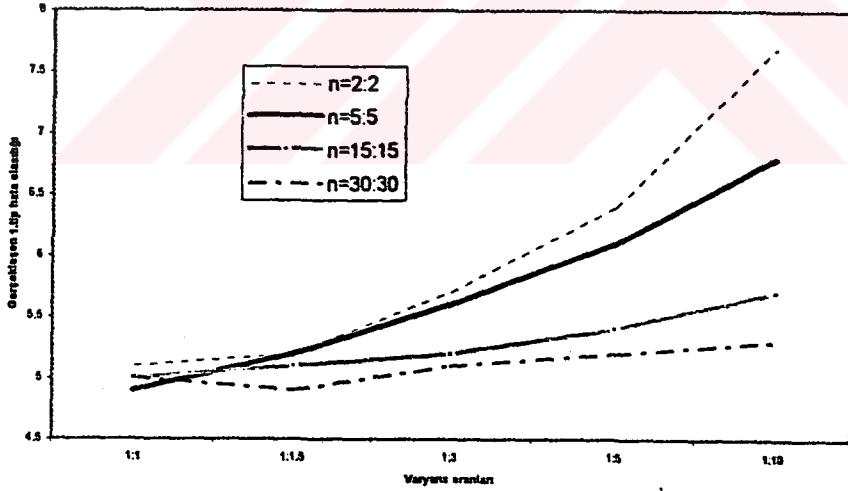
(S _i :S _j) Varyans oranları	(n _i :n _j) Örnek genişliği	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
1:1	2:2	5139	5.1
	5:5	4930	4.9
	15:15	4977	5.0
	30:30	5005	5.0
1:1.5	2:2	5189	5.2
	5:5	5192	5.2
	15:15	5054	5.1
	30:30	4940	4.9
1:3	2:2	5735	5.7
	5:5	5555	5.6
	15:15	5233	5.2
	30:30	6134	5.1
1:5	2:2	6430	6.4
	5:5	6104	6.1
	15:15	5401	5.4
	30:30	5220	5.2
1:10	2:2	7745	7.7
	5:5	6845	6.8
	15:15	5700	5.7
	30:30	5314	5.3

Varyans oranları 1:10 olduğunda ise gerçekleşen 1. tip hata olasılıkları başlangıçta kararlaştırılan % 5.0'den oldukça fazla olarak bulunmuştur. Çizelge 4.1'den de görüleceği üzere, örnek genişlikleri 2, 5, 15 ve 30 olduğunda 1. tip hata olasılıkları % 7.7, % 6.8, % 5.7 ve % 5.3 olarak gerçekleşmektedir.

Gruplardaki varyans oranlarının 1:5, 1:3, 1:5 ve 1:10 olduğu durumlarda 2, 5, 15 ve 30 örnek genişlikleri için türetilen veriler, yaygın olarak kullanılan logaritmik, ters ve karekök transformasyonlarına tabi tutulduktan sonra bütün

analizler yeniden yapılmış ve gerçekleşen 1. tip hata olasılıkları saptanmıştır. Bu durumda gerçekleşen 1. tip hata olasılıkları orijinal verilerdekiler gibi bulunmuşlardır. Bu sonuçlara göre, dağılım şeklinin normal, iki grup karşılaştırılmasında, örnek genişliklerinin eşit ve varyans oranlarının 1.5, 1:3, 1:5 ve 1:10 olduğu durumlarda gerçekleşen 1. tip hata olasılıklarında meydana gelen değişimlerin logaritmik, ters ve karekök transformasyonlarla engellenemeyeceği, ancak örnek genişliğini artırarak gerçekleşen α 'daki değişimin olumlu yönde azaltılabileceği söylenebilir.

Karşılaştırılan gruplardaki varyans oranları değiştikçe gerçekleşen 1. tip hata olasılıkları, örnek genişliklerine göre şekil 4.1'de verilmiştir. Şekil 4.1'den de görüleceği üzere, varyans oranları arttıkça, gerçekleşen 1. tip hata olasılıkları, başlangıçta kararlaştırılmış olan % 5.0'den büyük olmakta ve α 'daki bu büyüme örnek genişliklerinin nisbeten küçük olduğu durumlarda çok daha belirginleşmektedir.



Şekil 4.1. Grup sayısı 2 ve gruplardaki gözlem sayıları eşit olduğunda 100000 deneme sonucunda gerçekleşen I. tip hata ihtimalleri

4.1.2. İki grup ve gruplardaki gözlem sayıları birbirinden farklı iken elde edilen sonuçları

Bu koşuldaki verilerin varyans analizi tekniği ile analizinden sonra gerçekleşen 1. tip hata olasılıkları saptanmış ve hesaplama sonuçları Çizelge 4.2'de topluca verilmiştir.

Çizelge 4.2. Grup sayısı 2 ve gruplardaki gözlem sayıları birbirinden farklı olduğunda 100.000 deneme sonucunda red edilen hipotez sayıları

(S_i, S_j) Varyans oranları	(n_i, n_j) Örnek genişliği	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orjinal veri (X)	
		Sayı	%
1:1	2:5	5030	5.0
	2:10	5049	5.1
	2:20	4950	5.0
	5:10	5086	5.1
	5:20	4868	4.9
	10:30	4990	5.0
1:1.5	2:5	3687	3.7
	2:10	2763	2.8
	2:20	2154	2.2
	5:10	3713	3.8
	5:20	2814	2.8
	10:30	3083	3.1
1:3	2:5	2347	2.3
	2:10	887	0.9
	2:20	378	0.4
	5:10	2458	2.5
	5:20	925	0.9
	10:30	1340	1.3
1:5	2:5	1889	1.9
	2:10	450	0.4
	2:20	86	0.09
	5:10	1933	1.9
	5:20	426	0.4
	10:30	456	0.5
1:10	2:5	1538	1.5
	2:10	201	0.2
	2:20	5	0.005
	5:10	1592	1.5
	5:20	175	0.2
	10:30	356	0.4

Çizelge 4.2'den de görüleceği üzere varyanslar homojen olduklarında yapılan varyans analizi sonunda gerçekleşen 1. tip hata olasılıkları gruplardaki farklı gözlem sayılarından etkilenmemekte ve başlangıçta karşılaştırılan % 5.0 civarında kalmaktadırlar.

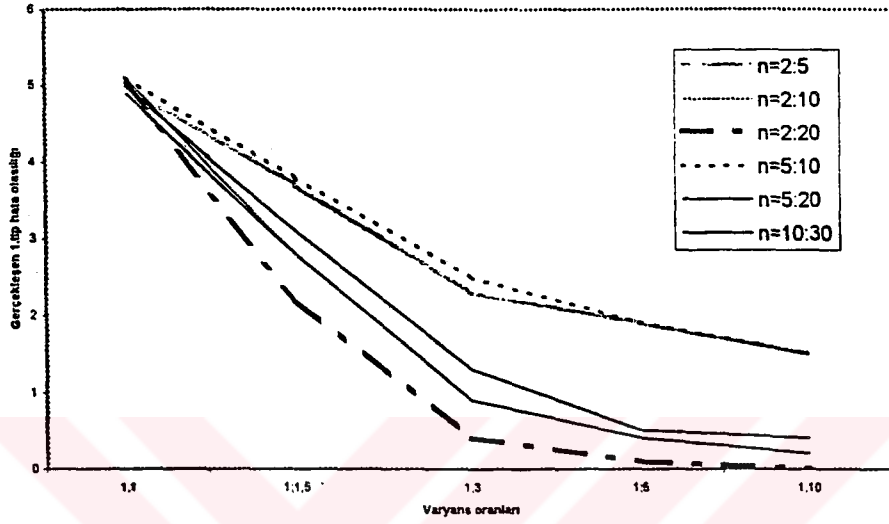
Gruplardaki varyans oranları 1:1.5 olduğunda ise gerçekleşen 1. tip hata olasılıkları başlangıçta karşılaştırılmış olan % 5.0'in altına düşmekte ve bu düşüş gruplardan birinde 2, diğerinde ise, sırasıyla 5, 10 ve 20 gözlem bulunduğunda, 1. tip hata olasılıkları % 3.7, % 2.8 ve % 2.2 olarak gerçekleşmektedir. Benzer eğilim gruplardan birinde 5, diğerinde ise, sırasıyla 10 ve 20 gözlem bulunduğunda da gözlenmektedir. Burada 1. tip hata olasılıkları % 3.8 ve % 2.8 olarak gerçekleşmektedir. Gruplardaki gözlem adetleri 10 ve 30 olduğunda ise gerçekleşen 1. tip hata ihtimali bir miktar yükselmekte fakat % 3.1 ile başlangıçta karşılaştırmış olan % 5.0'in oldukça altında kalmaktadır.

Çizelge 4.2'nin incelenmesinden varyans oranlarının 1:3, 1:5 ve 1:10 olduğunda da gerçekleşen 1. tip hata olasılıkları bakımından yukarıdakine benzer bir durum olduğu anlaşılmaktadır. Ne varki varyans oranları değiştikçe gerçekleşen 1. tip hata olasılıkları da belirgin bir şekilde azalmakta ve bu azalış gruplardaki gözlem sayılarının farklılaşmasıyla daha da güçlenmektedir. Varyans oranları 1:3 olduğunda gruplardan birinde 2 diğerinde 5 gözlem olduğunda gerçekleşen 1. tip hata olasılığı % 2.3 olmakda, gruplardan birinde 2 diğerinde ise 20 gözlem olduğunda ise adı geçen olasılık % 0.4'e inmektedir. Varyans oranları 1:10 olduğunda ise yukarıdaki olasılıklar % 1.5 ve % 0.005 olmaktadır.

Veriler logaritmik, ters ve karekök transformasyonlarına tabi tutulduktan sonra bütün analizler yeniden yapılmıştır. Yapılan hesaplamalar sonunda bu transformasyonların gerçekleşen 1. tip hata olasılığındaki düşüslere etkili olmadıkları sonucuna varılmıştır.

Gruplardaki varyans oranları ile gerçekleşen 1. tip hata olasılıkları, gruplardaki gözlem adetlerine göre şekil 4.2'de verilmiştir. Şekil 4.2'den de

görülebileceği üzere varyansların oranı değiştikçe gerçekleşen 1. tip hata olasılıkları başlangıçta kararlaştırılan % 5.0'in altına düşmekte ve bu düşüş gruplardaki gözlem sayıları dengesizleştikçe hızlanmaktadır. Elde edilen sonuçlar Şekil 4.2'de sunulmuştur.



Şekil 4.2. Grup sayısı 2 ve gruplardaki gözlem sayıları birbirinden farklı olduğundan 100000 deneme sonucunda ret edilen hipotez sayılarının dağılımı

4.1.3. Üç grup ve gruplardaki gözlem sayıları birbirine eşit iken elde edilen sonuçlar

Bu koşuldaki verilerin varyans analizi tekniği ile analizden sonra gerçekleşen 1. tip hata olasılıkları saptanmış ve hesaplama sonuçları Çizelge 4.3'te topluca verilmiştir.

Çizelge 4.3'ten de görüleceği üzere, gruplardaki varyanslar homojen olduklarında ve gruplardaki gözlemler birbirine eşit iken yapılan varyans analizleri sonucunda gerçekleşen 1. tip hata olasılıkları, örnek genişliklerinin 2, 5, 15 ve 30 olduğu durumlarda sırasıyla % 5.1, % 4.9, % 5.0 ve % 5.0 olarak, başlangıçta kararlaştırılan % 5 civarında kalmışlardır.

Gruplardaki varyans oranları 1:1.5:2 olduğu zaman verilerin analizi sonucunda gerçekleşen 1. tip hata olasılıkları varyans oranlarının 1:1:1 olduğu durumdakinden pek farklı değildir. Bu olasılıklar örnek genişliklerine göre % 5.3, % 5.3, % 5.1 ve % 5.2 olarak gerçekleşmekte olup, bunların başlangıçta kararlaştırılmış olan % 5.0'den pek farklı olmadığı görülmektedir.

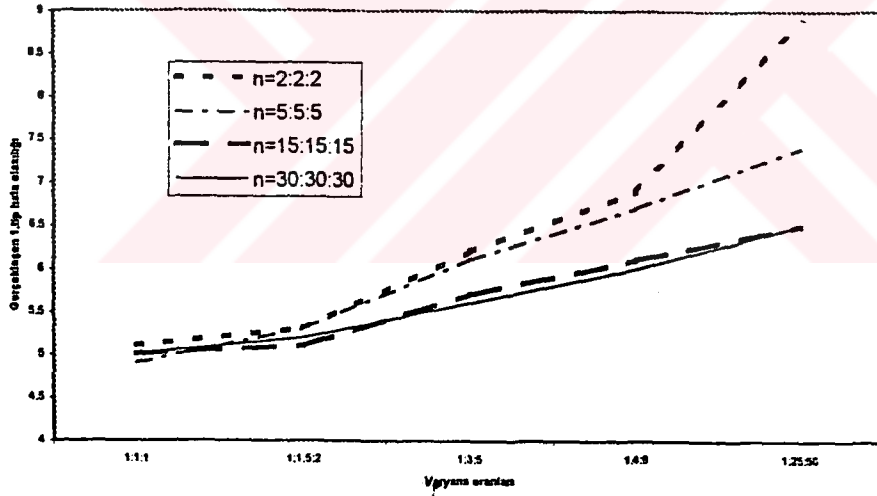
Çizelge 4.3. Grup sayısı 3 ve gruplardaki gözlem sayıları eşit olduğunda 100.000 deneme sonucunda red edilen hipotez sayıları

$(S_1:S_2:S_3)$ Varyans oranları	$(n_1:n_2:n_3)$ Örnek genişliği	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
1:1:1	2:2:2	5144	5.1
	5:5:5	4940	4.9
	15:15:15	4967	5.0
	30:30:30	5033	5.0
1:1.5:2	2:2:2	5276	5.3
	5:5:5	5323	5.3
	15:15:15	5142	5.1
	30:30:30	5201	5.2
1:3:5	2:2:2	6185	6.2
	5:5:5	6132	6.1
	15:15:15	5692	5.7
	30:30:30	5613	5.6
1:4:8	2:2:2	6877	6.9
	5:5:5	6655	6.7
	15:15:15	6107	6.1
	30:30:30	5969	6.0
1:25:50	2:2:2	8911	8.9
	5:5:5	7449	7.4
	15:15:15	6512	6.5
	30:30:30	6458	6.5

Varyans oranları 1:3:5, örnek genişlikleri de 2, 5, 15 ve 30 olduğu durumlarda gerçekleşen 1. tip hata olasılıklarında artış gözlenmiş, fakat bu artışın örnek genişliği arttıkça azalan bir artış olduğu saptanmıştır. Gruplardaki gözlem sayıları 5, 15 ve 30 olduğu durumlarda gerçekleşen 1. tip hata olasılıkları da sırasıyla % 6.2, % 6.1, % 5.7 ve % 5.6 olmuştur.

Gruplardaki varyans oranlarının 1:4:8 ve 1:25:50 olduğu durumlarda gerçekleşen 1. tip hata olasılıkları daha da artmış bu artış grupların varyansları arasındaki farkın büyümesiyle paralel doğrultuda artış göstermiş, gruplardaki gözlem sayısının artışıyla ise azalan bir artış olduğu ortaya çıkmıştır. Bu olasılıklar gruplardaki gözlem sayısına bağlı olarak sırasıyla şöyle elde edilmiştir. % 6.9, % 6.7, % 6.1, % 6.0 ve % 8.9, % 7.4, % 6.5, % 6.5.

Orjinal gözlemler logaritma, karekök ve ters transformasyonlara tabi tutularak bütün analizler yeniden yapılmıştır. Yapılan hesaplamalar sonunda bu transformasyonlar orijinal veridekilerden elde edilmiş olan 1. tip hata olasılıkları hemen hemen hiç değiştirmedeği görülmüştür. Varyans oranlarına bağlı olarak gerçekleşen 1. tip hata olasılıkları özet halinde Şekil 4.3'te sunulmuştur.



Şekil 4.3. Grup sayısı 3 ve gruplardaki gözlem sayıları eşit olduğunda 100000 deneme sonucunda red edilen hipotez sayılarının dağılımı

4.1.4. Üç grup ve gruplardaki gözlemler birbirinden farklı iken elde edilen sonuçlar

Üç grupta, varyans oranları birbirine eşit ve farklı iken, gruplardaki gözlem sayılarının da birbirinden farklı olduğu koşullarda uygulanan varyans analizi tekniği sonucunda gerçekleşen 1. tip hata olasılıkları Çizelge 4.4'te topluca verilmiştir.

Çizelge 4.4. Grup sayısı 3 ve gruplardaki gözlem sayıları birbirinden farklı olduğunda 100.000 deneme sonucunda red edilen hipotez sayıları

$(S_n:S_i:S_k)$ Varyans oranları	$(n_i:n_j:n_k)$ Örnek genişliği	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
1:1:1	2:5:10	5036	5.0
	5:10:15	4983	5.0
	5:10:20	5034	5.0
	10:20:50	5060	5.1
1:1.5:2	2:5:10	2726	2.7
	5:10:15	3135	3.1
	5:10:20	2743	2.7
	10:20:50	2380	2.4
1:3:5	2:5:10	1697	1.7
	5:10:15	2373	2.4
	5:10:20	1756	1.8
	10:20:50	2159	2.2
1:4:8	2:5:10	1496	1.4
	5:10:15	2159	2.2
	5:10:20	1476	1.5
	10:20:50	1118	1.1

Çizelge 4.4 incelendiği zaman gruplardaki varyans oranları birbirine eşit iken gruplardaki gözlem sayılarının farklı olduğunda gerçekleşen 1. tip hata olasılıklarının başlangıçta belirlenen % 5 gibi oldukları görülmektedir.

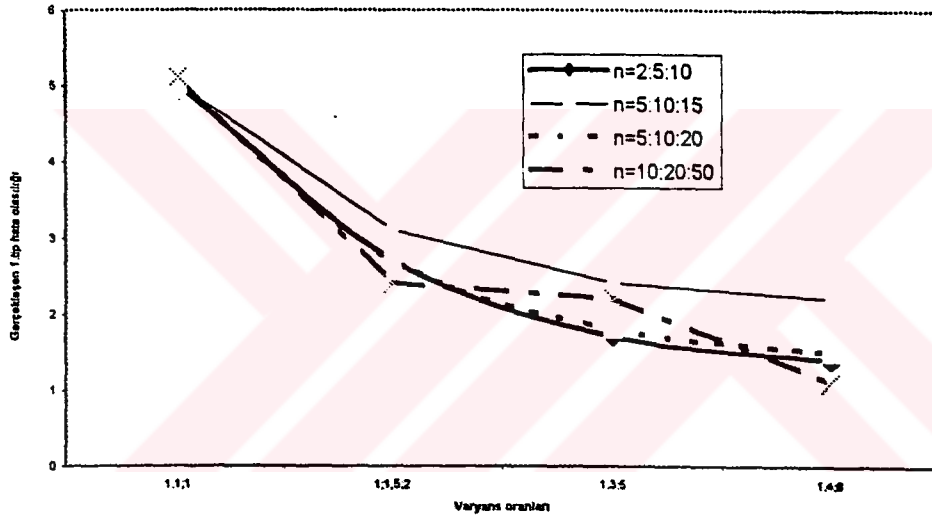
Gruplardaki varyans oranları 1:1.5:2 olduğu durumda gerçekleşen 1. tip hata olasılıkları farklı örnek genişlikleri için sırasıyla % 2.7, % 3.1, % 2.7 ve % 2.4 olarak elde edilmiştir. Bu durumda varyansların birbirlerinden farklılığı 1. tip hata olasılıklarının düşmesine sebep olmaktadır. Ayrıca gruplardaki örnek genişliklerinin birbirinden farkı arttıkça bu düşüş daha da belirginleşmektedir.

Varyans oranlarının 1:3:5 ve 1:4:8 olduğu durumlarda ise gerçekleşen 1. tip hata olasılıkları gruplardaki gözlem sayılarının birbirinden farklılığının büyümesiyle

daha da düşmektedir. Bu koşullarda elde edilen 1. tip hata olasılıkları sırasıyla % 1.7, % 2.4, % 1.8, % 2.2 ve % 1.4, % 2.2, % 1.5 ve % 1.1'dir.

Varyansların ve gruplardaki gözlem sayılarının birbirinden farklı olduğu durumlarda logaritmik, karekök ve ters transformasyonların gerçekleşen 1. tip hata olasılıkları üzerine etkisi araştırılmıştır. Transformasyonlar ile elde edilen sonuçların orijinal verilerde elde edilen sonuçlar ile çok yakın olduğu gözlenmiştir.

Gruplardaki varyans oranları değiştikçe gerçekleşen 1. tip hata olasılıkları Şekil 4.4'de verilmiştir.



Şekil 4.4. Üç grup gruplardaki gözlemler birbirinden farklı olduğunda 100000 deneme sonucunda red edilen hipotez sayılarının dağılımı

4.1.5. Dört grup ve gruplardaki gözlem sayıları birbirine eşit iken elde edilen sonuçlar

Bu koşullar altında elde edilen 1. tip hata olasılıkları Çizelge 4.5'te sunulmuştur.

Varyanslar eşit ve gruptaki gözlem sayılarında birbirine eşit iken gerçekleşen 1. tip hata olasılıkları başlangıçta belirlenen 1. tip hata olasılığı olan % 5 civarında kalmıştır. Yapılan simülasyon çalışmasında gruptaki gözlem sayıları 2, 5, 15 ve 30 olarak ele alındığında gerçekleşen 1. tip hata olasılıkları sırasıyla % 5.1, % 4.9, % 4.9 ve % 5.1 olarak elde edilmiştir. Bu durumda gruptaki gözlem sayıları ister küçük, ister büyük olsun gerçekleşen 1. tip hata olasılığında önemli bir artma veya azalma gözlenmediği ortaya çıkmıştır. Yani bu koşullar için örnek genişliğini artırmanın α üzerine önemli bir etkisi olmadığı söylenebilir.

Çizelge 4.5. Grup sayısı 4 ve gruptaki gözlem sayıları eşit olduğunda 100.000 deneme sonucunda red edilen hipotez sayıları

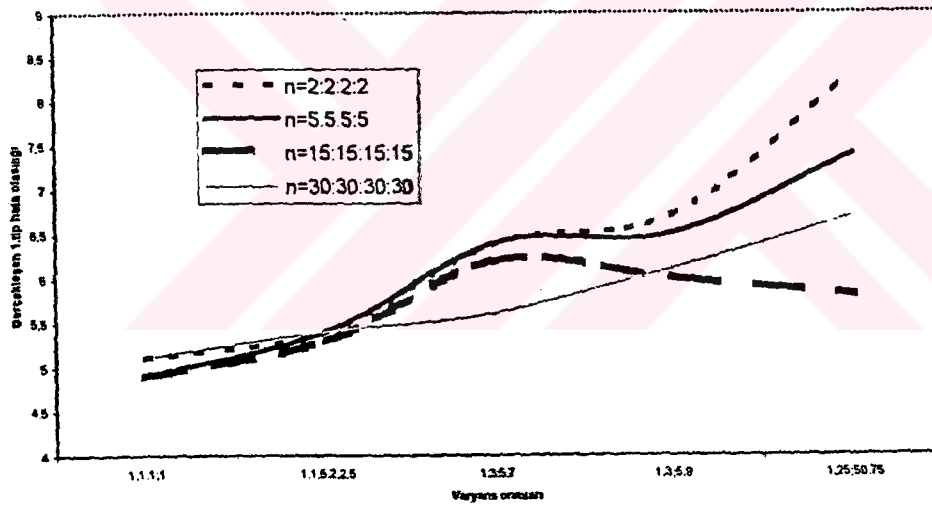
$(S_1:S_2:S_3:S_4)$ Varyans oranları	$(n_1:n_2:n_3:n_4)$ Örnek genişliği	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
1:1:1:1	2:2:2:2	5101	5.1
	5:5:5:5	4949	4.9
	15:15:15:15	4922	4.9
	30:30:30:30	5142	5.1
1:1.5:2:2.5	2:2:2:2	5434	5.4
	5:5:5:5	5445	5.4
	15:15:15:15	5343	5.3
	30:30:30:30	5401	5.4
1:3:5:7	2:2:2:2	6442	6.4
	5:5:5:5	6397	6.4
	15:15:15:15	6226	6.2
	30:30:30:30	5554	5.6
1:3:5:9	2:2:2:2	6692	6.7
	5:5:5:5	6488	6.5
	15:15:15:15	5956	6.0
	30:30:30:30	6111	6.1
1:25:50:75	2:2:2:2	8303	8.3
	5:5:5:5	7449	7.4
	15:15:15:15	6997	7.3
	30:30:30:30	6697	6.7

Gruptaki varyans oranları 1:1.5:2:2.5 olduğu durumda elde edilen 1. tip hata olasılıkları grupların örnek genişliği 2 iken % 5.4, 5 iken % 5.4, 15 iken % 5.3 ve 30 iken % 5.4 olarak tespit edilmiştir. Bu durumda varyans oranlarının değişmesinin gerçekleşen 1. tip hata olasılığını azda olsa bir miktar artırdığı söylenebilir.

Varyans oranları 1:3:5:9 ve 1:25:50:75 olduğunda gerçekleşen 1. tip hata olasılıklarının daha da arttığı, gruplardaki gözlem sayısındaki artışın da bir miktar azalmaya sebep olsa da gerçekleşen 1. tip hata olasılıklarının kararlaştırılan düzeyin oldukça üstüne çıkmasına sebep olduğu gözlenmektedir.

Bu koşullar için, logaritmik, ters ve karekök transformasyonlarından sonra hesaplamaların hepsi yeniden yapılmış, transformasyonların gerçeklerin 1. tip hata olasılıkları üzerine istenilen doğrultuda bir etki yapmadığı tespit edilmiştir.

Varyans oranlarına bağlı olarak, gerçekleşen 1. tip hata olasılıklarındaki değişimler Şekil 4.5'te özet halinde sunulmuştur.



Şekil 4.5. Grup sayısı 4 ve gruplardaki gözlem sayıları eşit olduğundan 100000 deneme sonucunda red edilen hipotez sayılarının dağılımı

4.1.6. Dört grup ve gruplardaki gözlem sayıları birbirinden farklı iken elde edilen sonuçlar

Varyanslardaki farklılıklar ve gruplardaki gözlem sayılarının birbirinden farklı olması durumunda gerçekleşen 1. tip hata olasılıkları Çizelge 4.6'da topluca sunulmuştur.

Gruplardaki varyans oranları birbirine eşit iken yani varyanslar homojen iken küçük ve büyük örnek genişliklerinde gerçekleşen 1. tip hata olasılıkları başlangıçta belirlenin % 5 civarında kalmaktadır. Bu durumda normal dağılım gösteren verilerde varyansların homojenlik önşartıda sağlanıyor ise örnek genişliklerinin ve gruplardaki gözlem sayılarının farklılığının gerçekleşen 1. tip hata olasılıklarını etkilemediği söylenebilir.

Gruplardaki varyans oranları 1:1.5:2:2.5 olduğunda farklı örnek genişlikleri için gerçekleşen 1. tip hata olasılıkları oldukça düşük olarak sırasıyla % 2.9, % 2.8, % 2.0 ve % 2.8 olarak elde edilmiştir.

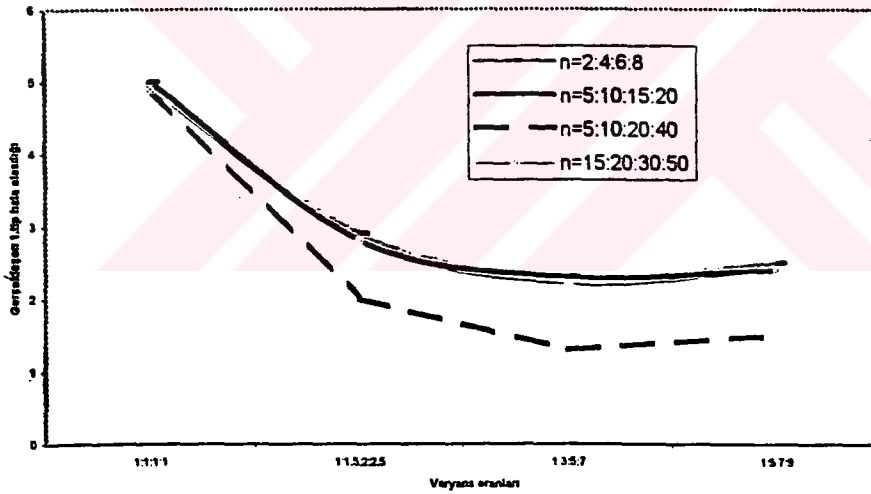
Çizelge 4.6. Grup sayısı 4 ve gruplardaki gözlem sayıları birbirinden farklı olduğunda 100.000 deneme sonucunda red edilen hipotez sayıları

(S ₁ :S ₂ :S ₃ :S ₄) Varyans oranları	(n ₁ :n ₂ :n ₃ :n ₄) Örnek genişliği	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orjinal veri (X)	
		Sayı	%
1:1:1:1	2:4:6:8	5020	5.0
	5:10:15:20	4980	5.0
	5:10:20:40	4947	4.9
	15:20:30:50	4938	4.9
1:1.5:2:2.5	2:4:6:8	2895	2.9
	5:10:15:20	2839	2.8
	5:10:20:40	1995	2.0
	15:20:30:50	2781	2.8
1:3:5:7	2:4:6:8	2314	2.3
	5:10:15:20	2261	2.3
	5:10:20:40	1251	1.3
	15:20:30:50	2181	2.2
1:5:7:9	2:4:6:8	2455	2.5
	5:10:15:20	2409	2.4
	5:10:20:40	1544	1.5
	15:20:30:50	2448	2.4

Varyans oranları 1:3:5:7 olduğunda ise bu olasılıklar % 2.3, % 2.3, % 1.3 ve % 2.2; varyans oranları 1:5:7:9 olduğunda % 2.5, % 2.4, % 1.5 ve % 2.4'tür. Bu durumda gruptaki varyansların farklılıkları arttıkça ve örnek genişlikleri arasındaki farklılıklar büyüdükçe gerçekleşen 1. tip hata olasılıklarının başlangıçta belirlenen % 5'in oldukça altında kaldığı gözlenmektedir.

Bu koşullarda veriler, logaritmik, ters ve karakteristik transformasyonlarına tabi tutulduktan sonra bütün analizler yeniden yapılmıştır. Yapılan hesaplamalar sonunda bu transformasyonların orijinal verilerde gerçekleşen 1. tip hata olasılıklarını değiştirmedikleri saptanmıştır.

Varyans oranlarına bağlı olarak gerçekleşen 1. tip hata olasılıklarındaki değişik Şekil 4.6'da sunulmuştur.



Şekil 4.6. Grup sayısı 4 ve gruptaki gözlem sayıları birbirinden farklı olduğunda 100000 deneme sonucunda red edilen hipotez sayılarının dağılımı

4.2. χ^2 Dağılımı Gösteren Veriler

1, 3, 5 ve 15 serbestlik dereceli χ^2 dağılımlarından alınmış tesadüf sayılarının, eşit ve farklı sayılarda gözlem içeren iki, üç ve dört gruba ayrılarak varyans analizine tabi tutulması sonucunda aşağıdaki sonuçlar elde edilmiştir.

4.2.1. Bir serbestlik dereceli χ^2 dağılımından herbirinde eşit sayıda gözlem bulunan gruplara ilişkin sonuçlar

Grup sayısının 2, 3 ve 4 olduğu durumda gruplardaki gözlem sayıları eşit iken gerçekleşen 1.tip hata olasılıklarına ilişkin sonuçlar Çizelge 4.7 de verilmiştir.

Çizelge 4.7. Bir serbestlik dereceli χ^2 dağılım gösteren veriler ve grup sayısı 2 ve gruplardaki gözlem sayısı eşit olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayıları

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:2	5041	5.0
	5:5	3184	3.2
	15:15	4265	4.3
	30:30	4692	4.7
3	2:2:2	5462	5.5
	5:5:5	3383	3.4
	15:15:15	3994	4.0
	30:30:30	4632	4.6
4	2:2:2:2	5719	5.7
	5:5:5:5	3602	3.6
	15:15:15:15	4198	4.2
	30:30:30:30	4584	4.6

Çizelge incelendiği zaman, iki grup olduğunda ve gruplardaki gözlem sayıları birbirine eşit iken gerçekleşen 1.tip hata olasılıklarının başlangıçta belirlenen % 5 düzeyini tam olarak korumadığı gözlenmektedir. Sadece herbir grupta iki gözlem var iken gerçekleşen α değeri ile başlangıçta belirlenen α değeri birbirine benzer çıkmıştır. Gruplardaki gözlem sayılarında gözlenen artış gerçekleşen α değerinin başlangıçta belirlenenden düşük çıkmasına sebep olmuştur. Grup sayısı iki ve gruplardaki gözlem sayıları 2, 5, 15 ve 30 olduğu durumda gerçekleşen α değerleri sırasıyla % 5, % 3.2, % 4.3 ve % 4.7 olarak bulunmuştur.

Grup sayısı 3 ve gruplardaki gözlem sayısı 2 iken gerçekleşen α değeri % 5.5 olarak bulunmuş, gözlem sayısı 5 iken gerçekleşen α değeri % 3.4'e düşmüş ve gözlem sayısı 15 ve 30'a çıktığı zaman ise gerçekleşen α değerleri % 4 ve % 4.6 olup biraz daha artmış fakat belirlenen seviyeye ulaşamamıştır.

Grup sayısının 4 olduğu koşullarda ise gerçekleşen α değerinde meydana gelen değişimler grup sayısının 3 olduğu koşullarla hemen hemen aynıdır. Burada da 1. tip hata olasılığı gruplardaki gözlem 2 olduğu durumda yükselmekte, % 5.7 iken gözlem adetleri 5 olduğunda düşmekte gözlem adetleri 15 ve 30 olduğunda ise biraz daha yükselmektedir. Gözlem sayısındaki artışlar gerçeklerin 1. tip hata olasılığını başlangıçta belirlemesi seviyeye yaklaştırmaktadır..

Bu tip verilere logaritmik, karekök ve ters transformasyonların uygulanması gerçekleşen α değerlerini istenilen düzeye getirmemiştir. Ancak logaritmik transformasyona tabi tutulan verilerde gerçekleşen 1. tip hata olasılıkları % 4.3 ile % 5.1 arasında değişmişlerdir. Karekök transformasyonunda ise bu değerler % 4.9 ile % 5.7 arasında değişmektedirler.

Ters (1/X) transformasyonu ile yapılan analizlerde ise gerçekleşen 1. tip hata olasılıkları % 0.6 ile % 3.2 arasında değişmektedir. Bu durumda ters (1/X) transformasyonunun uygun olmadığı söylenebilir.

4.2.2. Bir serbestlik dereceli χ^2 dağılımında herbirinde farklı sayıda gözlem bulunan gruplara ilişkin sonuçlar

Gruplardaki gözlem sayıları birbirinden farklı iken grup sayılarına bağlı olarak gerçekleşen 1. tip hata olasılıkları Çizelge 4.8'de topluca sunulmuştur.

Çizelge 4.8. Üç serbestlik dereceli χ^2 dağılımı gösteren 2, 3 ve 4 grupta ve gruplardaki gözlem sayılarının farklı iken 100.000 deneme sonucunda red edilen hipotez sayıları

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:5	4521	4.5
	2:10	5738	5.7
	2:20	5979	6.0
	5:10	3920	3.9
	5:20	4468	4.5
	10:30	4385	4.4
3	2:5:10	5299	5.3
	5:10:15	4372	4.4
	5:10:20	4540	4.5
	10:20:50	4744	4.7
4	2:4:6:8	4806	4.9
	5:10:15:20	4702	4.7
	5:10:20:40	5164	5.2
	15:20:30:50	4462	4.5

Grup sayısı 2 olduğunda ve gruplardaki gözlemler birbirinden farklı iken orjinal verilere uygulanan varyans analizi sonuçlarında gözlem 1. tip hata olasılıkları, örnek genişliklerine bağlı olarak başlangıçta belirlenen α değerinden düşük veya yüksek bulunmuştur. Örnek genişliklerinin oldukça farklı olduğu durumda (2 ve 20) gerçekleşen 1. tip hata olasılığının % 6'ya ulaştığı dikkati çekmektedir.

2 grup için gerçekleşen sonuçlara benzer sonuçlar 3 ve 4 grupta da elde edilmiştir.

Orijinal verilere logaritmik transformasyon uygulandıktan sonra varyans analizi tekniği kullanıldığında grup sayısı 2 ve gruplardaki gözlem sayıları da 2:5 iken gerçekleşen 1. tip hata olasılığı % 4.8 olup başlangıçta belirlenen % 5'e çok yakın bir değer olmuştur. Gruplardaki gözlem sayıları 2:10 iken gerçekleşen α değeri % 5.4, 2:20 iken % 5.0 ve 5:10 iken % 4.7 ve 10:30 iken % 4.8 olarak tespit edilmiştir. Bu sonuca göre gerçekleşen α değerindeki değişimlerin orijinal verilere göre daha az olduğu görülmektedir.

Grup sayısı 3 ve gruplardaki gözlem sayıları da 2:5:10, 5:10:15, 5:10:20 ve 10:20:50 iken gerçekleşen α değerleri sırasıyla % 5.1, % 4.7, % 4.8 ve % 5.0 olup 2 gruptaki sonuçlara benzer sonuçlar elde edilmiştir.

Grup sayısı 4'e çıkınca gerçekleşen α değerler gruplardaki gözlem sayılarına bağlı olarak sırasıyla % 5.9, % 4.9, % 5.1 ve % 4.7 olarak gerçekleşmişlerdir. Bu sonuçlara göre logaritmik transformasyonu sonunda gerçekleşen 1. tip hata olasılıklarının kararlaştırılan % 5 değerine nisbeten yakın bulundukları söylenebilir..

Karekök transformasyonunda elde edilen sonuçlar incelendiği zaman gerçekleşen 1. tip hata olasılıklarının % 4.4 ile % 5.0 arasında olduğu görülmüştür. Buna göre karekök transformasyonunun bu tip veriler için etkili sonuçlar verdiği ileri sürülebilir.

Ters (1/X) transformasyon sonucu elde edilen 1. tip hata olasılıkları ise 2 grupta % 1.8 ile % 11.9, 3 grupta % 2.2 ile % 9.2, 4 grupta ise % 2.3 ile % 7.8 arasında muntazam olmayan değişimler göstermişlerdir. Bu durumda Ters (1/X) transformasyonunun uygun olmadığı görülmektedir.

4.2.3. Üç serbestlik dereceli χ^2 dağılımında, herbirinde eşit sayıda gözlem bulunan gruplarda ilişkin sonuçlar

Üç serbestlik dereceli χ^2 dağılımından alınan veriler herbirinde eşit sayıda gözlem bulunan 2, 3 ve 4 gruba ayrıldıktan sonra uygulanan varyans analizi sonuçlarında elde edilen 1. tip hata olasılıkları çizelge 4.9'da sunulmuştur.

Çizelge 4.9. Üç serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı eşit olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayıları

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:2	5235	5.2
	5:5	4419	4.4
	15:15	4778	4.8
	30:30	4522	4.5
3	2:2:2	5392	5.4
	5:5:5	4395	4.4
	15:15:15	4561	4.6
	30:30:30	4846	4.8
4	2:2:2:2	9346	5.4
	5:5:5:5	4304	4.3
	15:15:15:15	4686	4.7
	30:30:30:30	4795	4.8

Çizelgenin incelenmesinden grup sayısı 2 ve gruplardaki gözlem sayılarının sırasıyla 2, 5, 15 ve 30 olduğu durumlarda gerçekleşen 1. tip hata olasılıklarının sırasıyla % 5.2, % 4.4, % 4.8 ve % 4.5 olduğu görülmektedir. Benzer sonuçlar grup sayısının 3 ve 4 olduğu durum içinde geçerlidir. Bu durumda gruplardaki gözlem sayısının artışının gerçekleşen α değerinin başlangıçta belirlenen düzeye yaklaşmasında etkili olduğunu fakat grup sayısındaki değişimin ise bunu etkilemediği söylenebilir.

Bu tip verilere logaritmik transformasyon uygulandığı zaman gerçekleşen α değeri grup sayısı ve gruplardaki gözlem sayılarına bağlı olarak % 4.7 ile % 5.0 arasında değişmekte olup, bu sonuçlar karşılaştırılan % 5'e orjinal verilerdekinden daha yakındırlar.

Karekök transformasyonu uygulanan verilerden elde edilen sonuçlarda ise gerçekleşen α değerleri % 4.1 ile % 5.3 arasında olup α değerindeki değişmeler, logaritmik transformasyonunkine nazaran biraz daha fazladır.

Ters transformasyonu uygulanması durumunda ise gerçekleşen α değerleri % 0.05 ile % 4.5 arasında olup oldukça düşükdürler.

Bu sonuçlara göre bu tip veriler için en uygun sonuçların logaritmik ve karekök transformasyonlar ile elde edildiği görülmektedir.

4.2.4. Üç serbestlik dereceli χ^2 dağılımında, herbirinde farklı sayıda gözlem bulunan gruplara ilişkin elde edilen sonuçlar

Bu koşullara ilişkin elde edilen sonuçlar Çizelge 4.10'da topluca sunulmuştur. Sonuçlar incelendiği zaman grup sayısı 2 ve gruplardaki gözlem sayıları 2:5, 2:10, 2:20, 5:10, 5:20 ve 10:30 iken gerçekleşen 1. tip hata olasılıklarının sırasıyla % 4.9, % 5, % 4.9, % 4.6, % 4.8 ve % 4.8 olduğu görülmektedir. Bu değerler karşılaştırılan % 5'e oldukça yakındırlar.

Grup sayısı 3 ve 4 olduğunda gerçekleşen 1. tip hata olasılıkları % 4.3 ile % 5.4 arasında değişmekte olup 2. gruba oldukça benzer sonuçlar elde edilmiştir. Bu durumda orjinal verilere uygulanan varyans analizi sonucunda gruplardaki gözlem sayılarındaki değişmelerin ve grup sayısındaki değişmelerin gerçekleşen α değerini pek etkilemediği, fakat başlangıçta karşılaştırılan % 5 seviyesinden biraz düşük olduğu söylenebilir.

Çizelge 4.10. Üç serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı farkı olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayıları

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:5	4948	4.9
	2:10	5002	5.0
	2:20	4901	4.9
	5:10	4555	4.6
	5:20	4760	4.8
	10:30	4780	4.8
3	2:5:10	4998	5.4
	5:10:15	4641	4.3
	5:10:20	4771	4.7
	10:20:50	4874	4.8
4	2:4:6	4929	5.0
	5:10:15:20	4849	4.9
	5:10:20:40	4773	4.7
	15:20:30:40	4707	4.7

Bu tip verilere logaritmik transformasyon uygulandığı zaman gerçekleşen 1. tip hata olasılıkları % 4.8 ile % 5.1 arasında değişmiştir. Benzer durum karekök transformasyonu uygulandığında da gözlenmiştir.

Ancak ters ($1/X$) transformasyon uygulandığında, farklı durumlarda gerçekleşen α değeri bazen oldukça düşük (% 3.2) bazende oldukça yüksek (% 8) olarak bulunmuştur.

4.2.5. Beş serbestlik dereceli χ^2 dağılımında, herbirinde eşit sayıda gözlem

bulunan gruplara ilişkin elde edilen sonuçlar

Herbirinde eşit sayıda gözlem bulunan, 2, 3 ve 4 grup halinde düzenlenen verilere ilişkin elde edilen sonuçlar 4.11'de verilmiştir.

Çizelge 4.11. Beş serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı eşit olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayısı

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:2	5128	5.1
	5:5	4551	4.6
	15:15	4884	4.9
	30:30	4555	4.6
3	2:2:2	5227	5.2
	5:5:5	4611	4.6
	15:15:15	4703	4.7
	30:30:30	5053	5.1
4	2:2:2:2	5262	5.3
	5:5:5:5	4609	4.6
	15:15:15:15	4642	4.6
	30:30:30:30	5009	5.0

Orjinal verilerde grup sayısı 2 ve gruplardaki gözlem sayıları 2, 5, 15 ve 30 iken gerçekleşen 1. tip hata olasılıkları sırasıyla % 5.1, % 4.6, % 4.9 ve % 4.6 olarak bulunmuştur. Grup sayısı 3 ve 4 iken elde edilen sonuçlar grup sayısı 2 olduğu durumlardaki gibidir.

Bu tip verilere logaritmik transformasyon uygulandıktan sonra yapılan varyans analizleri ile gerçekleşen 1. tip hata olasılıkları % 4.7 ile % 5.1 arasında değişmiştir.

Karekök transformasyonu uygulandıktan sonra gerçekleşen 1. tip hata olasılıkları % 4.8 ile % 5.1 arasında bulunmuştur.

Ters transformasyon uygulanan verilerden elde edilen sonuçlarda gerçekleşen 1. tip hata olasılıkları başlangıçta belirlenen % 5 seviyesinden daha düşük bulunmuştur.

4.2.6. Beş serbestlik dereceli χ^2 dağılımında, birbirinde farklı sayıda gözlem bulunan gruplara ilişkin sonuçlar

Gruplardaki gözlem sayılarının birbirinden farklı olduğu durumlarda elde edilen sonuçlar, Çizelge 4.12’de verilmiştir.

Çizelge 4.12. Beş serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı farklı olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayısı

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:5	4936	4.9
	2:10	5009	5.0
	2:20	4743	4.7
	5:10	4540	4.5
	5:20	4689	4.7
	10:30	4907	4.9
3	2:5:10	5059	5.1
	5:10:15	4702	4.7
	5:10:20	4848	4.8
	10:20:50	4914	4.9
4	2:4:6:8	4884	4.9
	5:10:15:20	4917	4.9
	5:10:20:40	4816	4.8
	15:20:30:50	4771	4.8

Grup sayısı 2 ve gruplardaki gözlem sayıları 2:5, 2:10, 2:20, 5:10, 5:20 ve 10:30 iken gerçekleşen 1. tip hata olasılıkları sırasıyla % 4.9, % 5, % 4.7, % 4.5, % 4.7 ve % 4.9 olarak bulunmuştur. Grup sayısı 3 ve 4 iken gerçekleşen 1. tip hata olasılıkları ise % 4.7 ile % 5.1 arasında olup başlangıçta belirlenen % 5 seviyesine biraz daha yaklaşan değerler almışlardır.

Bu tip verilere logaritmik ve karekök transformasyonları uygulandıktan sonra yapılan varyans analizleri sonucunda gerçekleşen 1. tip hata olasılıkları % 4.8 ile

% 5.1 arasında değişmektedir. Bu durumda da verilere logaritmik ve karek ek transformasyonları uygulamanın gerçekleşen 1. tip hata olasılıkları üzerine daha etkili olacağı gözlenmektedir.

Ters (1/X) transformasyonu uygulandıktan sonra elde edilen 1. tip hata olasılıklarının % 3.6 ile % 6.6 arasında değiştiği gözlenmiştir. Bu durumda ters transformasyonun gerçekleşen 1. tip hata olasılıklarını başlangıçta belirlenen seviyeye yaklaştırmada etkili olmadığı söylenebilir.

4.2.7. Onbeş serbestlik dereceli χ^2 dağılımında, herbirinde eşit sayıda gözlem bulunan gruplara ilişkin sonuçlar

Bu koşullar altında yapılan varyans analizlerine göre elde edilen 1. tip hata olasılıkları Çizelge 4.13'te sunulmuştur.

Çizelge 4.13. Onbeş serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı eşit olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayısı

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:2	5099	5.1
	5:5	4845	4.8
	15:15	4908	4.9
	30:30	4674	4.7
3	2:2:2	5149	5.1
	5:5:5	4867	4.9
	15:15:15	4936	4.9
	30:30:30	5211	5.2
4	2:2:2:2	5210	5.2
	5:5:5:5	4779	4.8
	15:15:15:15	4960	5.0
	30:30:30:30	5199	5.2

Grup sayısı 2 ve gruptaki gözlem sayıları 2, 4, 15 ve 30 iken gerçekleşen 1. tip hata olasılıkları sırasıyla % 5.1, % 4.8, % 4.9 ve % 4.7 olarak tespit edilmiştir.

Grup sayısı 3 ve 4 iken gerçekleşen 1. tip hata olasılıkları % 4.8 ile % 5.2 arasında değişmekte olup bunların başlangıçta belirlenen % 5 seviyesine yakın değerler aldığı gözlenmiştir. Bu durumda serbestlik derecesinin artışının gerçekleşen α değerinin başlangıçta belirlenen seviyeye yaklaşmasını sağladığı söylenebilir.

Bu tip verilere logaritmik ve karekök transformasyonları uygulandıktan sonra elde edilen 1. tip hata olasılıkları % 4.9 ile % 5.2 arasında değişmekte olup orjinal verilerden daha iyi sonuçlar vermiştir. Böylece orjinal verilere logaritmik ve karekök transformasyonu uygulamanın daha etkili sonuçlar vereceği söylenebilir.

Ters ($1/X$) transformasyonu uygulandıktan sonra gerçekleşen 1. tip hata olasılıkları % 4.2 ile % 5.1 arasında değişmekte olup genellikle % 5'in altında kalmışlardır.

Serbestlik derecesi arttıkça ters transformasyon sonucunda gerçekleşen 1. tip hata olasılıkları başlangıçta belirlenen % 5 seviyesine yaklaşmaktadır. Ancak orjinal, logaritmik ve karekök transformasyonlarına nazaran α 'daki değişimlerin daha fazla olduğu gözlenmiştir.

4.2.8. Onbeş serbestlik dereceli χ^2 dağılımında, herbirinde farklı sayıda gözlem bulunan gruplara ilişkin sonuçlar

Bu koşullar altında elde edilen 1. tip hata olasılıkları Çizelge 4.14'te sunulmuştur.

Çizelge 4.14. Onbeş serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruptaki gözlem sayısı farklı olduğu durumda 100.000 deneme onucunda red edilen hipotez sayısı

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:5	4872	5.0
	2:10	5048	5.0
	2:20	4777	4.8
	5:10	4879	4.9
	5:20	4911	4.9
	10:30	4854	4.9
3	2:5:10	5001	5.0
	5:10:15	5088	5.1
	5:10:20	4840	4.8
	10:20:50	5075	5.1
4	2:4:6:8	4834	4.8
	5:10:15:20	5070	5.1
	5:10:20:40	5055	5.1
	15:20:30:50	5059	5.1

Çizelge incelendiği zaman grup sayısı 2, 3 ve 4 iken ve gruptaki gözlem sayıları birbirinden farklı iken gerçekleşen 1. tip hata olasılıklarının % 4.8 ile % 5.1 arasında değişmekte oldukları ve başlangıçta belirlenen % 5 seviyesine oldukça yakın değerler aldıkları gözlenmektedir. Gruplardaki gözlem sayılarının farklı olması gerçekleşen 1. tip hata olasılıkları üzerinde olumsuz bir etki yapmamıştır.

Verilere logaritmik ve karekök transformasyonları uygulandıktan sonra elde edilen 1. tip hata olasılıkları da % 4.8 ile % 5.1 arasında olup orijinal verilere benzer sonuçlar elde edilmiştir.

Ters transformasyon (1/X) sonucunda ise α değerleri % 4.5 ile % 5.3 arasında olup diğer durumlara nazaran biraz daha iyi sonuçlar elde edilmiştir. Burada da gerçekleşen 1. tip hata olasılıkları genellikle % 5'in altında kalmışlardır.

Bütün bu açıklamalar altında χ dağılımı gösteren veriler için serbestlik derecesindeki artışının, gerçekleşen 1. tip hata olasılıklarının başlangıçta belirlenen seviyeye yaklaşmasını sağladığı görülmektedir. Bu durumun sebebi ise serbestlik dereceleri büyüdükçe χ^2 dağılımlarının şekillerinin simetrikleşmeye başlayıp, normal dağılımı andırmasıdır. Serbestlik derecesinin düşük olduğu durumlarda ise logaritmik ve karekök transformasyonlarının nisbeten etkili sonuçlar verdiği söylenebilir.

4.3. Genel Değerlendirmeler

Yapılan simulasyon çalışmaları sonunda normal dağılımı gösteren verilerde iki grup karşılaştırmasında gruptaki gözlemler eşit olduğunda varyans oranları 1:3 olana kadar varyansların heterojenliği 1 tip hata olasılıklarını etkilememektedir. Heterojenlik bozuldukça yani varyansların oranları 1:3, 1:5, 1:10 olduğunda gerçekleşen 1. tip hata olasılığı da, özellikle küçük örneklerde hızla büyümekte ve % 7.7'ye kadar varmaktadır. Büyük örnekler ise varyansların heterojenliğini oldukça iyi tolere etmektedirler. Bu durumda varyansların oranı 1:10 bile olsa gözlem adetleri 30 olduğunda 1. tip hata olasılığı % 5.3 olarak gerçekleşmektedir. Gruptaki gözlem adetleri farklı olduğunda ise varyansların oranının 1:1.5 olmasından itibaren gerçekleşen 1. tip hata olasılığı hızla küçülmekte (% 3.7) ve varyansların oranı 1:10 olduğunda % 0.005'e kadar inmektedir. Bu küçülmede gruptaki gözlem adetlerinin farklılığı önemli bir rol oynamaktadır. Uygulamada yaygın olarak kullanılan log, $\sqrt{X}$ ve $1/X$ transformasyonları yukarıda değinilen durumları değiştirmemektedir.

Üç grup karşılaştırmasında da yukarıdakilere benzer sonuçlar elde edilmiştir. Gruptaki gözlem adetleri eşit olduğunda varyans oranlarının 1:1.5:2 olması gerçekleşen α 'yı etkilememekte, gruptaki gözlem adetleri farklı olduğunda ise bu varyans oranlarında gözlem adetleri eşit (2'şer gözlem) varyans oranları 1:4:8 olduğunda α % 6.9 olarak gerçekleşmekte, gözlem adetleri farklı (örneğin; 2:5:10)

olduğunda ise gerçekleşen 1 tip hata olasılığı % 1.4'e inmektedir. Uygulanan transformasyonlar da gerçekleşen 1. tip hata olasılıklarını pek değiştirmemişlerdir.

Dört grup karşılaştırmasında da benzer eğilimler gözlenmiştir. Varyans oranları 1:1:1:1 yerine 1:1.5:2:2.5 olduğunda gruptaki gözlemler eşit olduğuna α % 5.4 civarında gerçekleşmekte, gruptaki gözlem adetleri farklılaştığında (örneğin; 2:4:6:8) ise gerçekleşen α % 2.9'a inmektedir.

Dikkat edileceği üzere varyanslar homojen olmadığında ve de gruptaki gözlem adetleri farklı olduğunda gerçekleşen 1. tip hata olasılığı önceden kararlaştırılmış olan % 5'in oldukça altına düşmektedir. Bunun sebebi, bu çalışmada büyük varyanslı populasyonlardan alınmış olan örnek genişliklerinin küçük varyanslı populasyonlardan alınmış olanlardan daha büyük olmasıdır. Varyansı büyük olan populasyonlardan daha dar, varyansı küçük olan populasyonlardan daha geniş örnekler alınmış olsa idi gerçekleşecek 1. tip hata olasılığının % 5'den fazla olabileceğine dikkat edilmelidir.

Bu durumda, dağılımının şekli normal olduğunda varyansların heterojenliğinin örnek genişliğinin artırarak bir miktar dengelenebileceği söylenebilir. Ne varki, mesela, dört grup karşılaştırmasında varyansların oranı; 1:25:50:75 olduğunda gruptaki gözlem adetleri 30'ar olsa bile 1. tip hata olasılığının % 6.7 olarak gerçekleştiği göz önünde tutulmalıdır. Uygulamada yaygın olarak kullanılan log, $\sqrt{X}$ ve $1/X$ transformasyonları ise bütün düzenlerde gerçekleşen 1. tip hata olasılıklarını etkilememişlerdir.

1. serbestlik dereceli χ^2 dağılımı gösteren verilerle yapılan analizlerde 2, 3, 4 grup karşılaştırmasında gruptaki gözlem adetleri eşit ve de farklı olduklarında gerçekleşen 1. tip hata olasılıkları genellikle % 5'e oldukça yakın gerçekleşmekte, logaritmik ve $\sqrt{X}$ transformasyonları gerçekleşen 1. tip hata olasılıklarını % 5'e oldukça yaklaştırmaktadır. 1. tip hata olasılıklarını dengelemek bakımından ters ($1/X$) transformasyonunun uygun olmadığı sonucuna varılmıştır.

3, 5 ve 15 serbestlik dereceli χ^2 dağılımlarında da benzer durumlar gözlenmektedir. Özellikle 15 serbestlik dereceli χ^2 dağılımında 2, 3 ve 4 grup karşılaştırmasında, gruplardaki gözlemler eşit ve farklı olduğunda 1. tip hata olasılıklarının % 4.7 ile % 5.2 arasında gerçekleştikleri dikkati çekmektedir. Aynı serbestlik dereceli χ^2 dağılımlarından yapılan grup karşılaştırmalarında grupların varyansları homojendir. Dolayısıyla varyansların homojenliği ön şartı yerine gelmiş olmaktadır. Serbestlik dereceleri büyüdükçe χ^2 dağılımlarının şekli simetrikleşmeye dolayısıyla da normal dağılımı andırmaya başladığından gerçekleşen 1. tip hata olasılıkları da % 5'e yaklaşmaktadırlar. Buradan da varyans analizi tekniğinin ön şartlarından olan "Varyansların homojenliği" ön şartının 1. tip hata olasılığının oluşumunda diğer önemli ön şart olan "normal dağılım" ön şartından daha belirleyici olduğu görülmektedir.



KAYNAKLAR

- ATKINSON, A.C., 1973 Testing transformations to normality. Journal of the Royal Statistical Society, Series B, 35, 473-479.
- AVRALIOĞLU, Z., 1976. İstatistik. Ankara.
- BEK, Y.L, EFE, E., 1995. Araştırma ve Deneme Metotları I. Çukurova Üniversitesi Ziraat Fakültesi Yayınları. No: 71, Ders Kitabı, Adana.
- BICKEL, P.J. and DOKSUM, K.A., 1982. Analysis of transformation revisited. J.A.S.A, 76, 296-311.
- BOX, G.E.P. and COX, D.R., 1982. Analysis of transformation revisited, rebutted. J.A.S.A, 77, 209-210.
- BOX, G.E.P. and COX, D.R., 1964. An analysis of transformations, Journal of the Royal Statistical Society B 26, 211-252.
- CARROLL, R.J., 1980. A robust method for testing transformations to achieve approxiamete normality. Journal of the Royal Statistical Society B. 42. 71-78.
- CHERKAOUİ, M., 1976. Transformation to Normality: Variations of the Power Transformation Parameters under various Sociological Condition. Quality an Quantity 10, (2), Netherlands.
- CANAVOS, G.C. and MILLER, D.M., 1993. An Introduction to Modern Business Statistics, California.
- ÇETİNÖZ, E., 1980. Box-cox Transformasyon Yönteminin Soya Fasulyesi Verilerine Uygulanması. Yüksek Lisans Tezi. E.Ü.E.H.B.E. Bornova.
- DRAPER, N.R. and D.R. COX. 1969. On distributions and their transformations to normality. Journal of the American Statistical Association, 77, 103-108.
- DOKSUM, K.A., 1984. Discussion of 'The analysis of transformed data' by Hinkley and Runger, J.A.S.A, 79, 316-319.
- DÜZGÜNEŞ, O., KESİCİ, T. ve GÜRBÜZ, F. 1983. İstatistik Metotları. A.Ü. Ziraat Fakültesi Yayınları: 861, Ders Kitabı, s. 1-229, Ankara.
- DÜZGÜNEŞ, O., KESİCİ, T., KAVUNCU, O., GÜRBÜZ, F., 1987. Araştırma ve Deneme Metotları (İstatistik Metodları-II). A.Ü. Ziraat Fakültesi Yayınları: 1201. Ders Kitabı: 295, Ankara.

- EMERSON, J.D. and STOTO, M.A., 1982. Exploratory methods for choosing power transformations. *Journal of the American Statistical Association*, 77, 103-108.
- ERGEN, M.Ö., 1979. Varyans analizinin dayandığı varsayımlardan bazılarının gerçekleşmediği durumlarda Box ve Cox transformasyon yönteminin kullanılması üzerine bir araştırma. Doktora tezi, E.Ü.E.H.B.E, Bornova.
- GAUDRY and LAFERRIERE, 1989. The Box-cox transformation, power invariance and a New Interpretation *Economic Letters*, 30(1), 27-29.
- HINKLEY, D.V., 1975. "On power Transformations to symmetry", *Biometrika*, 62, 101-112.
- HINKLEY, D.V. and RUNGER, G., 1984. The analysis of transformed data. *J.A.S.A*, 79, 280-283.
- HINZ, P. and EAGLES, H.A., 1976. Estimation of transformation for the analysis of some agronomic and genetic experiments. *Crop science*. 16, 280-283.
- HOYLE, M.H., 1973. Transformations an introduction and bibliography. *International Statistical Review*. 41, 203-223.
- KOTTLER, F., 1950. The distribution of portide sizes. *J. Franklin Inst.* 250, 339-356.
- MANLEY, B.F.J., 1976. Expenertial data transformations, *The statistician* 25: 37-42.
- RUBIN B. DONALD, 1984. Distinguishing Between the scale of the Estimand and the Transformation to Normality. *J.A.S.A Volum 79, Number 386, Therory and Methods section*.
- RUPPERT, D., 1989. Transformations to symmetry and homosceclasticity. *Jasa*. 406, 437-446.
- SAKIA, R.M., 1992. The Box-cox transformation techique, a review. *The Statistician* 169-178.
- SCHIESSELMAN, J.I., 1973. Data transformation in two-way analysis of variance, *Journal of the American Statistical Association*, 68, 369-378.
- SEZGİN, F., 1972. Varyans Analizinin Dayandığı Faraziyeler ve Bunların Bozulmasından Doğan Durumlar. Doktora Tezi, Atatürk Üniversitesi Ziraat Fakültesi. Erzurum.

- SMITH, J.E. KEITH., 1976. Data Transformations in Analysis of Variance. Journal of Verbal Learning and Verbal Behavior. 15, (3) 339-346, U.S.A.
- SPITZER, J.J., 1978. A Monte Carlo investigation of the Box-cox transformation in small samples J.A.S.A, 73: 488-495.
- SPITZER, J.J., 1982. A fast and Efficient Algorithm for Estimation of Parameters in Models with the Box and cox transformation J.A.S.A, 77(380), December 760-66.
- SPITZER, J.J., 1984. Variance Estimates in models with Box-cox transformation: Implications for Estimations and Hypothests Testing Review of Economics and Statistics; 66(4), 645-52.
- ŞENYURT, S., 1979. Varyans analizinin dayandığı varsayımların gerçekleştiği ve gerçekleşmediği durumlarda F testlerinin incelenmesi. Yüksek Lisans Tezi, E.Ü.E.H.B.E. Bornova.
- THÖNI, H., 1963b. Die nomographische Bestimmung des Logarithmischen Durehschnittes Experientia 19, 219-219.
- THÖNI, H., 1967. Transformations of variables used in the Analysis of Experimental and observational Data A. REVIEW Technical Report Number 7, Statistical Laboratory Iowa State university Ames. Iowa.
- TUKEY, J.W., 1955. On the Comparative Anatomy of Transformations. The Annals of Mathematical Statistics. 28, 602-632.
- WOOD, R. KEVIN, 1989. Variance reduction using nonlinear controls and transformations. Comm, Statist. Simulation. 18, 2, 655-672.
- ZEKİ, A. NECİP, 1979. Normallikten Sapmanın F Testine Etkisi Üzerinde Simulasyon Yöntemi İle Araştırmalar. Doktora Tezi. A.Ü. Ziraat Fakültesi Zirai Genetik ve İstatistik Kürsüsü.

ÖZGEÇMİŞ

1951 yılında Elazığ'da doğdu. İlk, orta ve lise öğrenimini Elazığ'da tamamladı. 1976 yılında AİTİA Ekonomi-Maliye Bölümü'nden mezun oldu.

Aynı yıl AİTİA İstatistik Enstitüsünde Yüksek Lisans öğrenimine başlayarak 1978 yılında yüksek lisans derecesini aldı. Halen Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesinde Öğretim Görevlisi olarak çalışmaktadır. Evli ve iki çocuk babasıdır.

